The Roles of Perfectionism, Rumination and Exam Stress in the Onset of Depressive Symptoms in Adolescence

Musil, Anna Sofia Fredrika

Awarding institution:
King's College London

The Roles of Perfectionism, Rumination and Exam Stress in the Onset of Depressive Symptoms in Adolescence

Anna Sofia Fredrika Musil

Thesis submitted in partial fulfilment of the degree of Doctorate in Clinical Psychology

Institute of Psychiatry, Psychology and Neuroscience

King's College London

May 2018

ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude to my supervisors Eleanor Leigh and Patrick Smith for their guidance and support through all stages of writing this thesis. Eleanor, you have been so generous with your time, and I have learnt so much from your advice and questions. Patrick, I have appreciated your availability to meet with me, and your help with problem-solving when most needed.

A huge thank you to all the young people who took part in my empirical project, and to Cath and Abbi. It would not have been possible without you. Eleanor Hearder, Phoebe Wheeler, Kenny Chiu and Clem Edwards; your help with administering the questionnaires was also invaluable.

I would like to thank May Elliott-Joshi for generously offering her time and help in quality assessing all studies included for the systematic review.

It has been a huge privilege to meet and work with so many different people over the three years on this course. I would especially like to thank the brave people who have consented to their stories being shared in the second volume of this thesis; Your individual strengths and resilience has really inspired me. I am also very grateful to the supervisors that I have had. I developed a lot on each placement.

On a personal note, I would like to thank my fellow trainees, but especially Maxine, Angeliki and May. I feel so lucky to have got to know you. I would also like to thank Julian, who not only was of such great help during the most stressful times, but who I have felt so cared for and supported by. I would also like to thank my Swedish friends, for their cheerleading throughout the years.

Finally, I would like to thank my family for their love, support and encouragement. I could not have done this without you.

Table of Contents

EMPIRICAL PROJECT

The Roles of Perfectionism, Rumination and Exam stress in the onset of Depressive Symptoms in Adolescence

Table of Contents

List of Table and Figures

Abstract

The present study explored perfectionism as a cognitive vulnerability of depression. A group of 135 adolescents, aged between 15 and 16 years old, completed measures of self-oriented and socially-prescribed perfectionism, rumination and depression, three weeks before an important exam period. Symptoms of depression were measured again four weeks later, after the exams had finished but before the results had been shared. A cross-sectional mediation analysis revealed that both self-oriented and socially prescribed perfectionism could predict depression and that these associations were fully mediated by rumination. After the exams, pupils on average reported an improvement in mood. Socially-prescribed perfectionism was however associated with higher scores of depression at Time 2 compared to their peers, which could not be explained by pre-exam rumination. This study adds to the existing literature suggesting the two types of perfectionism may have different developmental trajectories. Implication and advice for future research and clinical work are discussed.

1. Introduction

1.1. Adolescent Depression

Depression is one of the most common mental health problems experienced by young people globally (Polanczyk, Salum, Sugaya, Caye, & Rohde, 2015). In adults and adolescents alike, depression is characterised by a change in affect, including persistent feelings of sadness or anhedonia; which in younger populations may be expressed as irritability. Disturbed sleep, a loss of appetite, lack of concentration, and thoughts surrounding hopelessness or despair are also common (Kessler, Avenevoli & Merikangas, 2001).

Over a five-year period, between the ages of 13 and 18, the lifetime prevalence of diagnosable mood disorders surges from 8.4% to 15.4% (Merikangas et al., 2010). A comparable number of young people, in addition, report subclinical levels of depressive symptoms (Wesselhoeft, Sørensen, Heiervang & Bilenberg, 2013; Kessler et al., 2001). There is also considerable evidence of continuity through to adulthood. For example, in a prospective study on young adults, the annual rate of major depressive disorder was 3.7% among people who had not previously reported difficulties with their mental health, compared to 9% for those who had experienced symptoms of depression in adolescence (Lewinsohn, Rohde, Klein & Seeley, 1999). Furthermore, three quarters of adults with recurrent or chronic depression retrospectively reported their first episode to have started before the age of 18 (Kim-Cohen et al., 2003). Epidemiological studies have found that symptoms of depression can be observed in a comparable number of boys and girls during childhood. However, starting from mid-adolescence approximately twice as many young women as men report being affected by depression (e.g. Twenge & Nolen-Hoeksema, 2002; Hankin, Mermelsteain & Roesch, 2007; Costello, Copeland, & Angold, 2011; Hankin et al., 2015; Cyranowski, Frank & Young, 2000).

Depression can have a devastating impact across all dimensions of a young person's life, and has been associated with interpersonal difficulties, poorer academic achievements and increased risk of attempted suicide (Klein, Shankman, Lewinsohn, & Seeley, 2009). These risks appear to persist; depression in adolescence has been linked to enduring difficulties, including reduced global functioning and unemployment in early adulthood (Fergusson & Woodward, 2002).

The increased prevalence rates, emergence of gender differences, and long-term impact on quality of life have led adolescence to be called ‘a critical time’ to understand how depression develops (Abela & Hankin, 2008; Hankin et al., 1998; Avenevoli et al. 2015). Exploring the causes and mechanisms leading to these difficulties is an important clinical question, as it can help inform interventions. It could also help to identify young people who are at greater risk of experiencing difficulties with their mood so that support can be put in place earlier. It may further be used in the development of prevention programmes to be delivered universally for young people, for example, in schools (Ingram & Price, 2011; Gladstone, Beardslee & O’Connor, 2011).

1.2. **Vulnerability-Stress Models of Depression**

Stress is believed to be a precursor to depression, either following significant life changes such as bereavement or divorce (Hammen, 2005; Monroe & Hadjiyannakis, 2002), or through a build-up of daily hassles, for example arguments with friends or parents, or worries about school work (Sim, 2000). When asked, most people retrospectively report to have encountered a stressful life event within a month prior to noticing their first symptoms of depression (Hankin & Abela, 2005). The transition to puberty has furthermore been linked to an increase in aversive experiences (Ge, Conger & Elder, 2001; Stikkelbroek, Bodden, Kleinjan, Reijnders & van Baar, 2016). Hence, the risk of stress may be particularly relevant to adolescents.

Multi-wave, longitudinal studies (e.g. Hankin et al., 2007; O'Connor, Rasmussen, & Hawton, 2010) have demonstrated that negative events indeed do play a significant role in the onset of depressive symptoms. However, stressful life events alone are not enough to predict changes to affect, and the majority of people who encounter aversive experiences do not become depressed (Hankin, 2006). A main objective has therefore been to explain why, given similar circumstances, some young people experience a strong negative emotional reaction to events, whereas others remain unscathed (Hankin, Fraley and Abela, 2005). An influential way to explain these differences has been using the diathesis-stress model (Meehl, 1962; Monroe & Simons, 1991). It proposes that each person has a unique set of vulnerabilities, which precipitate and maintain depression through the interaction with life stressors.

One type of diathesis is cognitive vulnerabilities. Thought processes are hypothesised to relate to depressive symptoms, for example through the perception of, reactions to, and judgements made when confronted with stressors (e.g. Ingram, Miranda and Segal, 1998). A person's experience of depression will hence depend on the severity of their vulnerabilities, the qualities of the stressor, and the cognitions involved (Abela & Hankin, 2008). Broader frameworks have been proposed, including Beck's cognitive theory (Beck, 1967), the Response Style theory (Nolen-Hoeksema, 1991) and the Personality Predisposition theory (Blatt & Zuroff, 1992), which have been well researched in the adult literature (see Scher, Ingram & Segal, 2005, for review). Evidence for these theories has also been found among younger populations (e.g. Abela, Brozina, & Haigh, 2002; Hankin & Roesch, 2005; Lewinsohn, Joiner & Rohde, 2001).

The need for further studies on cognitive vulnerabilities amongst adolescents, and their roles in the aetiology of depression, has however been highlighted (Abela & Hankin, 2008). Firstly, because adolescence is an age associated with neuroanatomical change, linked to alterations in both affect and cognitions (Keshavan, Giedd, Lau, Lewis & Paus, 2014). This is, for example, the time in which thinking is believed to become more crystallised and trait-like (Ingram, 2003; Ingram & Luxton, 2005; Cole, Jacquez, & Maschman, 2001). These developmental considerations mean findings from one age group are harder to generalise across young people. Effect sizes of cognitive vulnerabilities have for example been found to be small for children, but become larger during adolescent years (Lakdawalla, Hankin, & Mermelstein, 2007). Comparatively less is also known about how specific vulnerabilities (e.g. inferential style or self-criticism) interact with particular stressors, either congruent or incongruent with the person's beliefs. This has been identified as an interesting area for further research, as findings may give indications of when and how vulnerabilities first emerge (Abela & Hankin, 2008).

1.3. **Perfectionism as a Cognitive Vulnerability to Depression**

One cognitive factor linked to depression is 'perfectionism', the striving towards high standards followed by critical self-evaluation and distress when there is a risk of the demands not being met (Frost, Marten, Lahart, and Rosenblate, 1990). Observed as a transdiagnostic vulnerability factor (Egan, Wade, & Shafran, 2011), perfectionism has also been associated with anxiety (e.g. Hewitt et al., 2002), eating disorders (Dour and

Theran, 2011; Nilsson, Sundbom, & Hägglöf, 2008), and increased risk of suicidal behaviours amongst adolescents (Bibeau and Dupuis, 2007; Boergers, Spirito, and Donaldson, 1998; Enns, Cox and Inayatulla, 2003).

Perfectionism is sometimes thought to encompass both adaptive and maladaptive qualities (e.g. Frost et al., 1990; Rice and Preusser, 2002; Hamachek, 1978). It has been proposed that aspects of perfectionism, at times, can be motivating, help people to stay goal focused, and yield feelings of pride and success when a task has been achieved (Bieling, Israeli, Smith & Antony, 2003). Difficulties are thought to emerge when performance is judged with persistent self-criticism (Frost et al., 1990), when standards are endlessly moved to unobtainable levels (e.g. Bieling, Israeli, Smith & Antony, 2004), or when there is a discrepancy between personal standards and actual performance (Accordino, Accordino & Slaney, 2000; Afshar et al., 2011). Others have argued that no form of perfectionism is adaptive (Blatt, 1995; Shafran, Cooper & Fairburn, 2002; Flett & Hewitt, 2006).

Two of the most commonly used measures of perfectionism with adults are the Frost Multidimensional Perfectionism Scale (Frost et al., 1990) and the Hewitt-Flett Multidimensional Perfectionism Scale (Hewitt, Flett, Turnbull-Donovan & Mikail, 1991). Both measures share a conceptualisation of perfectionism as trait-like, to consist of both personal and relational aspects, and that the different dimensions of perfectionism relate to psychopathy in different ways. Frost and colleagues proposed perfectionism to be made up of six factors, including 'Concern over Mistakes', 'Doubts About Actions', 'Parental Expectations', 'Parental Criticism', 'Personal Standards' and 'Organization'. Hewitt and Flett instead, more broadly suggested perfectionism can be due to expectations perceived to come from other people (Socially-Prescribed Perfectionism), demands placed on oneself (Self-Oriented Perfectionism), or demands placed on others (Other-Oriented Perfectionism). In a comparison of two measures, Frost et al. (1993) suggested that Concerns over Mistakes and Socially-Prescribed Perfectionism both were linked to depression, while in contrast, Personal Standards and Other-Oriented Perfectionism were associated with elevated mood.

The Child and Adolescent Perfectionism Scale (Flett, Hewitt, Boucher, Davidson & Munro, 2000) is one of the most commonly used measures to assess perfectionism in youth. Based on the Hewitt-Flett Multidimensional Perfectionism Scale, it assesses

Socially-Prescribed Perfectionism (SPP) and Self-Oriented Perfectionism (SOP). Cross sectional studies with young people have reported correlations between both SOP and SPP and symptoms of depression (e.g. Stornelli, Flett and Hewitt., 2009; Hewitt et al., 2002). Hewitt et al. (2002) found a direct association between SPP and depressive symptoms, but suggested SOP interacted with social or achievement stressors to predict depression. These associations have also been investigated prospectively. In the 'Lifestyle and Coping' study (O'Connor et al., 2010), Scottish adolescents (N = 515) reported acute life stress, anxiety, depression, perfectionism, and incidents of self-harm over a 6-month period. Life stress was the main predictor of an increase in depressive symptoms over time, but SPP also accounted for some of the variance (1.6%). The studies outlined above are part of a growing body of evidence linking perfectionism to depression, yet the mechanism is not well understood.

1.4. **A Proposed Link Between Perfectionism and Depression**

Another cognitive process implicated in depression is rumination (Olson & Kwon, 2008). Alongside distraction and problem-solving, rumination is one of the reactions to symptoms of depression proposed in Nolen-Hoeksema's (1991) Response Styles Theory. By fixating on the causes and consequences of experiencing low mood, the repetitive negative thinking that characterizes rumination is thought to exacerbate and maintain depressive symptoms (Nolen-Hoeksema, Morrow & Fredrickson, 1993). Experimental, cross-sectional and longitudinal studies have consistently demonstrated links between rumination and depression in adults (e.g. Kirkegaard-Thomsen, 2006). Rumination has also been found to reduce mood in adolescents, in experimental (Park, Goodyer & Teasdale, 2004) as well as prospective studies (e.g. Abela & Hankin, 2011; Burwell & Shirk, 2007). Girls exhibit a greater tendency to ruminate than do boys (Hankin & Abramson, 2002; Burwell & Shirk, 2007), which have been found to account for some of the gender differences of depression.

Perfectionism and rumination are thought to be different processes; while perfectionism is a trait, rumination is a response style. Yet, they share common features. Like rumination, perfectionism is associated with repetitive and preoccupied self-monitoring (Frost, 1990). Internal and external events are scanned to look for signs of discrepancies between current and desired states, leading to cognitive biases in the encoding and processing of information (Hewitt & Genest, 1990; Nolen-Hoeksema et al., 2008). Flett,

Madorsky, Hewitt and Heisel (2002) found an association between rumination and perfectionistic cognitions, which has since been replicated with other samples of undergraduate students (Harris, Pepper, & Maack, 2008; O'Connor, O'Connor & Marshall, 2007). They suggested rumination mediate the link between perfectionism and depression (Flett et al., 2002).

Olson & Kwon (2008) further postulated that stressors, for example achievement tasks, may lead to intrusive, performance-related thoughts which might strengthen the association between perfectionism and rumination. They asked undergraduate students to complete measures of perfectionism, rumination and symptoms of depression at Time 1. Four weeks later, depressive symptoms were measured again. The interaction between perfectionism and rumination could predict a small, but statistically significant increase in symptoms of depression at Time 2, after accounting for depression scores at Time 1. Flett and colleagues (Flett, Coulter, Hewitt & Nepon, 2011) carried out a similar study with Canadian school pupils from year 7 and 8 but using a cross-sectional design. SOP correlated with rumination, but not with problem-solving or distraction. Surprisingly, no significant association was found between rumination and SPP. The authors suggested this link might be strengthened over adolescent years as more stressors are encountered, in line with the diathesis-stress model.

1.5. **Achievement Tasks as Stressors**

It has been suggested that young perfectionists have heightened sensitivity to personal failure, and therefore find achievement tasks particularly challenging. Examinations are stressful events for most young people. When year 11 pupils are asked what they worry about, homework and exams often top the list (Kyriacou & Butcher, 1993; Owen-Yeates, 2005). A lot is at stake for pupils, parents and schools alike, and result provide a public label which could impact on self-identity and self-worth (Denscombe, 2000).

Most studies assessing the impact of perfectionism in the classroom have used a correlational design. When the Almost Perfect Scale–Revised (Slaney, Rice, Mobley, Trippi & Ashby, 1996) has been employed as a measure of perfectionism, pupils who set high standards for themselves have tended to report higher self-esteem, fewer symptoms of depression, and achieved higher grade point averages (GPA) (Accordino et al., 2000; Stoeber & Rambow, 2007; Roohafza, Talaei, Sadeghi, Mackie & Sarafzadegan, 2010).

Negative reactions to imperfection (Stoeber & Rambow, 2007; Roohafza et al., 2010) or discrepancies between their desired and actual academic achievement (Accordino et al., 2000) have however been linked to symptoms of depression. Similarly, both socially-prescribed and self-oriented perfectionism were linked to achievement amongst a small subset of pupils (those with high academic ability). These traits were however also associated with lower mood and anxiety (Stornelli et al., 2009). Together these studies indicate that some aspects of perfectionism may lead to a reduction in mood around periods of high academic stress, for example following exams, but limitations have included the reliance on retrospective reports.

Einstein, Lovibond, & Gaston (2000) reported elevated symptoms of depression amongst final year high-school students (age 15-24 years), compared to normative data, at 10 weeks and again 10 days before exams; especially amongst those who endorsed high traits of SPP. In contrast, SOP was not strongly linked to depression. To our knowledge, this is the only prospective study on school aged pupils to date which has explored whether exams could be a trigger for depressive symptoms amongst adolescents vulnerable to harsh self-evaluation. Whilst this study has captured a build-up in distress during exam periods, it is unclear what happens after the stressor has passed. We would expect negative re-evaluation and rumination to be triggered once the exams have been completed, rather than beforehand.

1.6. **Present study**

In summary, previous research has given a plausible account for how perfectionism and rumination could form a pathway to depression, although studies have been few and often limited to undergraduate students rather than school aged pupils. Questions therefore remain about the replicability of these findings in a younger population. Most studies have also been correlational in their design. More prospective studies are therefore needed, for example exploring how proposed associations evolve in the context of life stressors (Morris & Lomax, 2015).

The present study aimed to further assess the perfectionism-rumination-depression link in a mid-adolescent UK sample, by building on the study designs and findings of Olson and Kwon (2008) and Flett et al. (2011). Pupils were asked to report symptoms of

depression before and after completing their Year 11 mock GCSE-exams[1]. As well as increasing ecological validity, this paradigm had the methodological advantage of having a predefined, universal stressor (Einstein et al., 2000). Few studies have looked at how exams affect mood among school-aged pupils, and findings have been inconsistent. Some have reported that exams exacerbate symptoms of depression (Robinson, Alexander & Gradisar, 2009), while others have failed to capture this main effect (e.g. Locker & Cropley, 2004). Overall, a slight increase in mood was therefore expected for most pupils after the exam had passed. For the adolescents who reported high levels of perfectionism and rumination, we would however hypothesize that performance without any immediate feedback would trigger repetitive thoughts about how they did, and in particular if any mistakes had been made, potentially leading to lower mood.

The present study consisted of two parts. First, a cross-sectional design was used to examine interrelationships between perfectionism, rumination and symptoms of depression in a sample of British school pupils. The main hypotheses were:

Hypothesis 1: Positive associations would be found between:

a) Perfectionism (SOP and SPP) and depressive symptoms
b) Perfectionism (SOP and SPP) and rumination
c) Rumination and depressive symptoms

Hypothesis 2: Rumination would mediate the link between perfectionism and depressive symptoms

The second part of the study was exploratory in nature, and aimed to prospectively examine how these cognitive vulnerabilities related to symptoms of depression following the encounter of a predictable stressor (mock GCSE-exams). For this part, three more hypotheses were proposed:

Hypothesis 3: There would be an overall change in depressive symptoms between Time 1 (before exams) and Time 2 (after the exams), with pupils reporting less depressive symptoms at Time 2.

[1] GCSE stands for General Certificate of Secondary Education. In England, these exams take place at the end of year 10 or 11 and determine choices of Sixth Form Education/College. They are also used as a requirement for further education, and for some jobs or apprenticeships (Education System in the UK, n.d.)

Hypothesis 4: Perfectionism (SPP and SOP) would be associated with higher depression scores at Time 2

Hypothesis 5: Rumination would mediate the relationship between trait perfectionism and depressive symptoms at Time 2

2. Method

2.1. **Study design**

The study consisted of two stages of data collection. The first screening (Time 1) was administered three weeks before the mock GCSE exam period had started. Measures of depressive symptoms, perfectionism and rumination were collected. The second screening (Time 2) took place 7-14 days after pupils had finished their mock exams. At Time 2, only measures of depression were collected.

2.2. **Ethical approval**

Ethical approval was obtained from King's College London Psychiatry, Nursing and Midwifery Research Ethics Subcommittee (HR-15/16-3736) on 21st October 2016. Subsequent modifications were also granted approval before screening commenced.

2.3. **Power calculation**

An a priori power calculation was conducted to estimate the number of participants required for the study. Effect sizes were based on findings by Olson and Kwon (2008). In a sample of undergraduate students, they reported the standardized beta coefficients for the relationship between perfectionism and rumination as $\beta = 0.29$, and between and perfectionism and depression symptoms as $\beta = 0.30$. To perform a cross-sectional mediation analysis, with 80% power at an alpha level of 0.05, a sample of approximately 150 participants was indicated.

2.4. **Pilot Study**

Prior to recruiting schools for the intended study, data was collected from a group of Year 9 pupils (N = 120) at the time of an end-of-year exam. This allowed us to assess the feasibility of undertaking a study during an exam period, and to finalise the research protocol. From the pilot project, it was estimated that one in every three schools approached would agree to take part in the project, and that the proportion of pupils who completed the full measure at both data collection points would be approximately 50 %.

2.5. **Recruitment Procedure**

Six state-funded academy schools local to King's College London or from surrounding London boroughs were contacted with study information and invitations. Five of the schools expressed an interest to take part in the project, and one school was able to participate in the time frame.

The participating school was located in an inner South London borough. With 207 pupils enrolled in year 11, it was among one of the largest schools approached. Demographic information was retrieved from the school's latest Office for Standards in Education report (2014). It revealed that one third of pupils identified as White British, while a large proportion of the young people had heritage from Black and minority ethnic groups. These figures were consistent with data reported for the borough by local authorities (Greater London Authority, 2014) suggesting the sample was representative of the general population living in the area.

2.6. **Inclusion/Exclusion Criteria**

All Year 11 pupils were invited to participate in the study. Due to the comprehension and literacy requirements of the self-report measures, pupils known to have a learning disability, autism spectrum disorder or difficulties understanding written English were however excluded from the study.

2.7. **Consent Procedures**

As the study was seen to be of low risk, and data collection was intended universally across the year groups, an opt-out procedure was used for parental consent. This was because research into school studies using opt-in parental consent has shown a risk of sample bias and underrepresentation of those from minority groups, of low socioeconomic status, or with poor health (e.g. Chartier et al., 2008; Wolfenden, Kypri, Freund & Hodder, 2009; Unger et al., 2004).

Pupils were given verbal and written information about the project before deciding if they wanted to take part. Printed information letters were also shared with parents and carers. In order to maximize opportunity to opt-out, electronic copies were also shared via the schools' emailing systems at least two weeks prior to data collection. Parents and carers were informed they could withdraw their young person from the project by returning the opt-out form to the class teacher, or by contacting the research leads.

On the days of data collection, the pupils were reminded about the project and that it was voluntary to take part. They were informed that they could withdraw from the study at any stage. The young people were also reminded that their responses would be kept confidential, but that answers to the questionnaires which indicated risk of harm towards themselves or others, or that the pupil was experiencing difficulties with low mood, would be shared with the schools Child Protection Lead. Pupils were asked to read and sign a consent form before completing any measures.

2.8. **Procedure**

Questionnaire packs were administered during form time at the beginning of the day. Pupils who did not want to take part continued with their individual learning. A member of the research team was available in the class room during data collection to answer questions. After data collection, a risk screening was carried out and concerns were followed up by the school's Child Protection Lead.

2.9. **Participant Characteristics**

Of the 207 pupils enrolled in the year group, fifteen pupils opted out from the study. Furthermore, 42 pupils were absent at the time of Time 1 screening, or questionnaire packs were not administered due to the exclusion criteria being met. Data from another

15 pupils was disqualified as not all questionnaire forms had been completed. This left a final sample of 135 pupils (response rate = 65%). Participants were 71 females and 64 males. The mean age was 15 years and 7 months (SD = 0.39 years).

All pupils who had completed questionnaires at Time 1 were asked to take part in the follow up screening at Time 2. At this time, four pupils opted out, and 33 pupils were absent. The remaining sample (n=102) consisted of 47 males and 52 females. This group of pupils did not differ from the group of pupils who completed Time 1 measures only, in age, gender, or symptoms of depression.

2.10. **Measures**

The Mood and Feelings Questionnaire (MFQ; Costello & Angold 1987) is a self-report measure designed to assess symptoms of depression in young people aged 6-17 years. In the present study, the 33-item long version was used. The MFQ comprises a list of statements and participants are asked to rate how well each statement describes how they have been feeling or acting over the past two weeks, using a scale from 0 to 2 (0 = not true, 1= somewhat true, 2 = true). Sample questions included *'I felt miserable and unhappy'* or *'I didn't enjoy anything at all'*. Higher total scores are associated with greater symptom severity. The MFQ has shown good internal consistency with $\alpha \geq 0.91$ (Daviss et al., 2006; Sund, Larsson & Wichstrøm, 2001). High test-retest reliability has also been reported, after three weeks (r = 0.84; Sund et al., 2001) and one month (ICC = 0.80; Daviss et al., 2006). In the present study, the internal reliability was excellent (α = 0.94).

The Children's Response Styles Questionnaire (CRSQ; Abela et al., 2000) is a 25-item self-report measure. Using a Likert scale from 0 = "almost never" to 4 = "almost always", respondents are asked to indicate how likely they are to act in a particular way when feeling low. Sample items included '*when you feel sad, do you read a book'* or '*when you feel sad, do think about how alone you feel'*. In this study, an adapted 21-item version used, as proposed by Abela, Aydin and Auerbach (2007). Accordingly, answers were grouped into two categories: rumination (13 items) and distraction/problem solving (8 items). Scores on the rumination subscale have consistently been linked to symptoms of depression (e.g. Abela et al., 2001; Abela et al., 2007). The internal consistency and test-retest reliability over a 4-week period has been reported as α = 0.82 and r = 0.72 for rumination, and α = 0.79 and r = 0.71 for distraction/problem-solving (Abela et al., 2007).

In the current sample, the internal reliability was $\alpha = 0.91$ for rumination and $\alpha = 0.70$ for distraction/problem-solving.

The Child and Adolescent Perfectionism Scale **(**CAPS; Flett et al., 2000) is a 22-item self-report measure used to assess traits of perfectionism. Using a Likert scale from 1 = "False, not at all true of me" to 5 = "Very true of me", respondents are asked to rate the extent to which they agree with statements measuring socially-prescribed (e.g. *'my family expects me to be perfect'*), self-oriented perfectionism (e.g.*'I get mad at myself when I make a mistake'*). Flett and colleagues (2016) demonstrated adequate internal consistency across for SOP $\alpha = 0.85$ and SPP $\alpha = 0.81$. They also reported the test-retest reliability over one year as $r = 0.80$ for SOP, and $r = 0.70$ for SPP, in a sample of Year 6 pupils.

3. Results

3.1. **Data analysis**

Data analysis was undertaken with IBM SPSS Statistics Software. An alpha level of 0.05 was used for all statistical tests. A data analysis plan was initially approved by Dr Daniel Stahl, and the final analysis presented in the result section were discussed approved by Dr Cedric Ginestet, both independent statisticians from King's College London Biostatistics Department.

3.2. **Treatment of missing data**

Questionnaire forms were screened for missing items. When less than 6% of values were missing, they were replaced by the participant's mean scores calculated independently for each questionnaire subscale. If more than 6% of items were missing on any of the questionnaire forms, the full data set was removed. This led to the exclusion of 15 participants at Time 1 and four participants at Time 2, and mean replacement of <1% of the included data.

3.3. **Test of normality**

Data for each variable was screened for outliers. Normality of the variables was assessed using the Kolmogorov-Smirnov test, and visual interpretations of histograms, normal Q-Q plots and box plots (see Appendix 9). Variables which were not normally distributed were processed with square root transformation prior to the use of parametric tests. When group differences were explored, Levene's test was used to verify homogeneity of variance. When assumptions could not be met, degrees of freedom were adjusted accordingly.

3.4. **Cross-sectional analysis**

3.4.1 Descriptive data

The means and standard deviations for each measure collected at Time 1 are presented in Table 1. Scores on the perfectionism measure (CAPS) and Response Style Questionnaire (CRSQ) were both comparable to values expected, based on previous studies (e.g. Flett et al., 2016; Flett et al., 2011; Abela et al., 2007). Symptoms of depression at Time 1, as reported by the Mood and Feelings Questionnaire (MFQ), were on average somewhat higher than anticipated. In a community sample of children and adolescents, a mean score of 11.6 was reported, and scores ≥ 29 were indicative of the clinical threshold for major depressive episode (Daviss et al., 2006). In the present sample, 18% of pupils reported scored above this clinical cut off, which may have accounted for a slight positive skew of the data.

Table 1. Descriptive Statistics (Means and standard deviations)

		All (N = 135)	**Females (N = 71)**	**Males (N = 64)**
Measure	**Range**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
CAPS- SOP	14-60	35.3 (9.5)	34.9 (10.0)	35.8 (9.0)
CAPS -SPP	11-50	25.5 (8.4)	25.7 (9.0)	25.4 (7.6)
CRSQ-Rum	13-51	25.1 (9.1)	26.7 (8.4)	23.3 (9.6)
CRSQ-D/PS	8-26	15.5 (4.2)	15.4 (4.3)	15.5 (4.2)
MFQ Time 1	0-58	16.7 (12.6)	20.4 (11.2)	12.7 (12.9)
MFQ Time 2	0-58	12.9 (11.8)	16.7 (11.6)	8.8 (10.8)

Note Abbreviations: CAPS (Child-Adolescent Perfectionism Scale), CAPS-SOP (Child-Adolescent Perfectionism Scale Self-Oriented Perfectionism subscale), CAPS-SPP (Child-Adolescent Perfectionism Scale Socially Prescribed Perfectionism subscale), CRSQ-Rum (Children's Response Styles Questionnaire Rumination subscale), CRSQ-D/PS (Children's Response Styles Questionnaire Distraction and Problem-Solving subscale)

A series of independent t-tests was carried out to compare the mean scores of boys and girls on each of the measures. Scores on the rumination subscale were higher for girls (M = 5.11, SD = 0.83) than for boys (M = 4.73, SD = 0.94), $t(126) = 2.44$, $p = 0.02$. Female participants also reported greater symptoms of depression (M = 4.33, SD = 1.31) than male participants (M = 3.07, SD = 1.81), $t(133) = 4.65$, $p < 0.01$. No other statistically significant differences were found, suggesting boys and girls reported similar levels of self-oriented and socially-prescribed perfectionism on the CAPS measure, and distraction/problem-solving on the CRSQ measure.

3.4.2. Hypothesis 1: Positive associations would be found between:

a) Perfectionism (SOP and SPP) and depressive symptoms

b) Perfectionism (SOP and SPP) and rumination

c) Rumination and depressive symptoms

A Pearson's product-moment correlation was used to assess relationships between the variables. As predicted, a strong, positive association was found between rumination and depressive symptoms ($r = 0.75$, $p < 0.01$). Significant but weak correlations were also found between perfectionism and depression, for SOP ($r = 0.17$, $p = 0.05$) and SPP ($r = 0.19$, $p = 0.03$). Rumination was also found to correlate with both SOP ($r = 0.28$, $p < 0.01$) and SPP ($r = 0.20$, $p = 0.02$).

Table 2. Correlations Between Dimensions of Perfectionism, Rumination, Distraction/ Problem-Solving, and Depressive Symptoms

	CAPS- SOP	CAPS -SPP^{T}	CRSQ-Rum^{T}	CRSQ-D/PS	MFQ Time 1^{T}
CAPS- SOP	1	0.48**	0.28**	0.17	0.17*
CAPS -SPP^{T}		1	0.20*	0.03	0.19*
CRSQ-Rum^{T}			1	0.13	0.75**
CRSQ-D/PS				1	0.02
MFQ Time 1^{T}					1

^{T} = Transformed data
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

3.4.3. Hypothesis 2: Rumination would mediate the link between perfectionism and depressive symptoms

To further explore the relationship between perfectionism and depressive symptoms, we employed a simple mediation model as outlined by Hayes (2013). The PROCESS macro for SPSS (model 4) was used to test whether perfectionism indirectly influenced symptoms of depression through its effect on rumination (see Figure 1.). The bootstrapping method, based on 5,000 bootstrap samples, were used to estimate bias-corrected confidence intervals. Separate analyses were run for self-oriented and socially-prescribed dimensions of perfectionism. We also ran each analysis twice to assess if the use of transformed data would make a difference to the overall findings. As they did not, the results for analysis using untransformed data have been presented below to ease the interpretation of effect sizes.

The first step of analysis was to establish a total effect of perfectionism (*X*) predicting symptoms of depression (*Y*) (path c). As can be seen in Table 3 and Table 4, this relationship was found both when SOP (b = 0.29, t(133) = 2.55, $p = 0.01$; $R^2 = 0.05$) and SPP (b = 0.30, t(133) = 2.36, $p = 0.02$; $R^2 = 0.04$) were used as the predicting variable. We then explored if perfectionism could predict rumination (*M*) (path a). Again, statistically significant links were found both between SOP and rumination (a = 0.27, $p < 0.01$), and SPP and rumination (a = 0.21, $p < 0.01$). Rumination could moreover predict symptoms of depression (path b), in models including SOP and SPP alike (b = 1.07 and b = 1.05 respectively, $p < 0.01$). These findings suggested that Baron and Kenny's assumptions of mediation (Baron & Kenny, 1986) had been met.

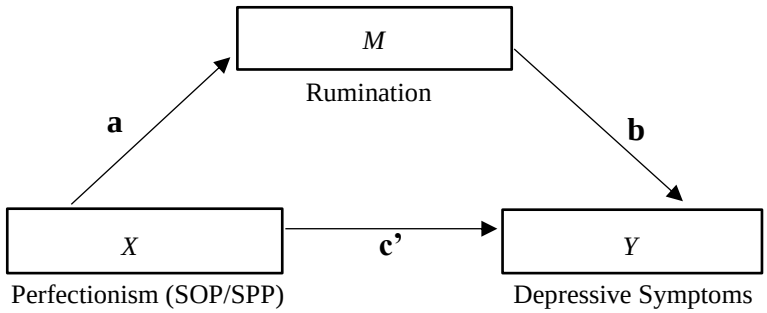


Figure 1. Simple mediation model

Bootstrapped confidence intervals were used to further estimate the indirect effects of perfectionism on depression. These were above zero both for SOP (ab = 0.29; CI 0.10 – 0.49) and SPP (ab = 0.22; CI 0.003 – 0.44) supporting the hypothesis of a predictive pathway, in which perfectionism leads to an increase in rumination, which in turn intensifies symptoms of depression. In contrast, the direct effect of perfectionism on depressive symptoms (c') were no longer statistically significant, for SOP (c' = -0.004, p = 0.96) or for SPP (c' = 0.08, p = 0.32), after controlling for rumination. Together, these findings suggest the links between perfectionism and rumination were fully mediated by rumination, for SOP and SPP alike.

Table 3. Mediation model coefficients (self-oriented perfectionism)

		Consequent						
		M (Rumination)				*Y* (Depressive Symptoms)		
Antecedent		Coeff.	*SE*	*p*		Coeff.	*SE*	*p*
X (SOP)	a	0.27	0.08	0.001	c'	-0.004	0.08	0.95
M (Rumination)		-	-	-	b	1.07	0.08	<0.01
		R^2 = 0.08				R^2 = 0.59		
		F (1,133) = 11.71, p < 0.01				F (2,132) = 98.11, p < 0.01		

Table 4. Mediation model coefficients (socially-prescribed perfectionism)

		Consequent						
		M (Rumination)				*Y* (Depressive Symptoms)		
Antecedent		Coeff.	*SE*	*p*		Coeff.	*SE*	*p*
X (SPP)	a	0.21	0.09	0.026	c'	0.08	0.09	0.33
M (Rumination)		-	-	-	b	1.05	0.08	<0.001
		R^2 = 0.04				R^2 = 0.60		
		F (1,133) = 5.07, p = 0.03				F (2,132) = 97.29, p < 0.01		

3.5. **Exploratory analysis**

The second part of the study was exploratory. It included a subset of pupils who had completed measures at Time 1 who also repeated the depression measure at Time 2.

3.5.1. Data preparation

Data was prepared as outlined in section 3.1. Three outliers were identified, whose changes in depression scores were three times smaller or larger than the interquartile (IQ) range. They were removed to improve the normality of the data prior to analysis. This left a sample of 99 pupils.

3.5.2. Descriptive statistics

To assess for group differences between pupils who participated at both time points compared to Time 1 only, a series of independent t-tests were carried out comparing mean scores of perfectionisms (SOP and SPP), rumination, distraction/problem-solving and depression scores at Time 1. No statistically significant differences in mean scores were found on any of the assessment measures. A Chi-square test was used to compare the proportion of male and female participant at Time 1 and Time 2. No statistically significant differences were found in gender distribution.

3.5.3. Hypothesis 3: There would be an overall change in depressive symptoms between Time 1 (before exams) and Time 2 (after the exam) with pupils reporting less symptoms of depression at Time 2.

A paired sample t-test was performed to compare mean depression scores between Time 1 and Time 2. On average, pupils reported more symptoms of depression four weeks before exams (M = 15.43, SD = 11.17) than they did in the immediate period after the exams had passed (M = 12.73, SD = 11.39), $t(98) = 4.12$, $p < 0.01$.

3.5.4. Hypothesis 4: Perfectionism (SPP and SOP) would be associated with higher depression scores at Time 2

A partial correlation was used to explore the relationship between perfectionism and symptoms of depression at Time 2, whilst controlling for symptoms of depression at Time 1. There was a modest, positive correlation between SPP (M = 24.69, SD = 7.45) and Time 2 depression scores (M = 12.73, SD = 11.39) whilst controlling for Time 1 depression score (M = 15.43, SD = 11.17), which was statistically significant (r = 0.23, p = 0.02). No statistically significant relationship was found between SOP (M = 34.76, SD = 8.90) and Time 2 depression scores, after controlling for symptoms of depression at Time 1 (r < 0.01, p = 0.95).

3.5.6. Hypothesis 5: Rumination would mediate the relationship between perfectionism and depressive symptoms at Time 2

To assess if rumination mediated the relationship between perfectionism and depression scores at Time 2, a simple mediation (as described in section 3.4.3) was run. Perfectionism (SOP/SPP) was used as the predicting variable (X), Time 2 depression scores were used as the outcome variable (Y) and rumination was used as the mediator (M). Symptoms of depression at Time 1 were controlled for as a covariate (C) as can be seen in *Figure 2*.

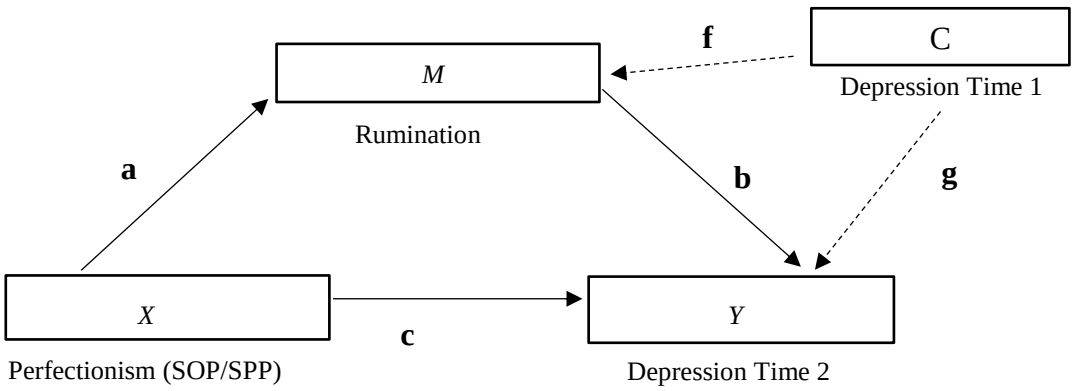


Figure 2. Simple mediation model with symptoms of depression at Time 1 as a covariate

Separate analyses were run for self-oriented and socially-prescribed perfectionism. After accounting for symptoms of depression at Time 1, the total effect of SOP on symptoms of depression at Time 2 was reduced (c = -0.001, p = 0.90). Furthermore, the indirect effect of SOP on depression with rumination as a mediator was not statistically different from zero (point estimate = 0.02, with a 95% bias-corrected bootstrap confidence interval of -0.01 to 0.06) suggesting self-oriented perfectionism could not predict symptoms of depression at Time 2.

In contrast, a total effect of SPP on symptoms of depression at Time 2 was found (c = 0.24, $p < 0.01$), even after accounting for symptoms of depression at Time 1. The indirect effect of SPP on depression with rumination as a mediator was not statistically different from zero (point estimate = 0.01, with a 95% bias-corrected bootstrap confidence interval of -0.01 to 0.05). These findings suggest that socially-prescribed perfectionism directly predicted symptoms of depression at Time 2, and rumination was not a prospective mediator.

Table 5. Mediation model coefficients (self-oriented perfectionism)

		Consequent						
		M (Rumination)				*Y* (Depressive Symptoms)		
Antecedent		Coeff.	*SE*	*p*		Coeff.	*SE*	*p*
X (SOP)	a	0.11	0.06	0.08	c’	-0.01	0.07	0.90
M (Rumination)		-	-	-	b	0.15	0.11	0.20
C (Depression Time 1)	f	0.52	0.05	<0.01	g	0.77	0.08	<0.01
		$R^2 = 0.52$				$R^2 = 0.59$		
		F (2,96) = 51.97, $p < 0.01$				F (3,95) = 98.11, $p < 0.01$		

Table 6. Mediation model coefficients (socially-prescribed perfectionism)

		Consequent						
		M (Rumination)				*Y* (Depressive Symptoms)		
Antecedent		Coeff.	*SE*	*p*		Coeff.	*SE*	*p*
X (SPP)	a	0.04	0.08	0.62	c’	0.23	0.08	< 0.01
M (Rumination)		-	-	-	b	0.13	0.11	0.23
C (Depression Time 1)	f	0.52	0.05	< 0.01	g	0.83	0.06	< 0.01
		$R^2 = 0.51$				$R^2 = 0.72$		
		F (2,96) = 49.17, $p < 0.01$				F (3,95) = 81.58, $p < 0.01$		

4. Discussion

The present study aimed to explore perfectionism as a possible cognitive vulnerability to depression in a mid-adolescent sample. The study had two aims; The first aim was to investigate if the link between perfectionism and depression was mediated by rumination as previously proposed by Flett and colleagues (2002). Secondly, to establish whether perfectionism could predict symptoms of depression prospectively, after the encounter of an academic stressor.

As an overview of the main findings, both self-oriented (SOP) and socially-prescribed (SPP) perfectionism were linked to symptoms of depression in a cross-sectional analysis. These associations were fully mediated by rumination. When symptoms of depression were measured again, five weeks later, socially-prescribed perfectionism reported at Time 1 was found to predict higher scores in depressive symptoms at Time 2. This effect was found to be direct. In contrast, self-oriented perfectionism was not associated with depression at Time 2. These findings suggest that young people react differently to achievement stressors, and that those who report socially-prescribed perfectionism may be at a higher risk of developing symptoms of depression in the context of such stress.

4.1. **The replication of a mediation model**

Building on the work by Flett et al. (2011) and Olson and Kwon (2008), the first aim of the study was to explore if perfectionism was associated with depression; and whether this relationship was direct or mediated by rumination. Measures were collected three weeks before the start of an important exam period. Pupils in the present sample were observed to report elevated symptoms of depression, compared to expected norms (Daviss et al., 2006). Bivariate correlations indicated associations between perfectionism and depression, perfectionism and rumination, and rumination and depression.

A simple mediation model (Hayes, 2013) was employed to further investigate the nature of these relationships. Both SOP and SPP were found to predict depressive symptoms. The two effects were fully mediated by rumination. These results were consistent with findings from an older sample of undergraduate students (Olson and Kwon, 2008). In contrast, Flett and colleagues (2011) did not find a significant association between SPP and rumination in a group of younger pupils (ages 12 to 14 years). They suggested that SPP and rumination may become more strongly paired as more stressors are encountered. Findings from the present study could be used in support of this hypothesis; during the transition from early to mid-adolescence, pupils would for example likely have noticed an increase in academic demands, which may have been accompanied by higher expectations from others. This may have led to a stronger internalisation of demands, and an increase in self-monitoring.

4.2. **Perfectionism as a prospective predictor of depression**

The second aim of the study was to explore if perfectionism could predict depression prospectively. Symptoms of depression were reassessed within two weeks of pupils finishing their exams, but before their results had been disseminated. Overall, the sample's mean score of depressive symptoms were significantly lower than prior to the exams. These findings may have reflected mood returning to baseline, or pupils experiencing elevated mood due to a sense of relief after the stressor had passed.

A partial correlation was used to explore the association between perfectionism at and symptoms of depression at Time 2, whilst controlling for Time 1 depression scores. Again, there was a modest, positive relationship between SPP and symptoms of

depression. Pupils endorsing this type of perfectionism was hence more likely than their peers to report higher levels of depression at Time 2; suggesting a slower recovery following the stressful event among pupils with this predisposition. This association was not found between SOP and depressive symptoms.

A mediation analysis was repeated, to further investigate if perfectionism and rumination measures collected at Time 1 could predict symptoms of depression at Time 2, and to explore the relationship between the variables. Symptoms of depression at Time 1 were also controlled for. As in the cross-sectional analysis, SPP was found to predict depression, but in this prospective model the effect was direct, and no longer mediated by rumination. SOP was not found to predict depression prospectively.

Some considerations may be offered to why these differences between SPP and SOP were found, which are not mutually exclusive. Firstly, it may be that SOP is more dynamic than SPP. This type of perfectionism may become more activated as a response to depression, rather than precipitating depressive symptoms.

Another explanation may be that SPP and SOP interact with distinct cognitive factors, which influences pupils' experiences of the exams. Young adults with high traits of SPP have for example been found more likely than their peers to believe that failures would have negative social consequences (Conroy, Kaye and Fifer, 2007). This may in turn activate catastrophic thinking related to the exams, which could impact on mood. In contrast, SOP has been linked to intrinsic motivation in achievement tasks, and to hope (Ashby, Dickinson, Gnilka & Nobel, 2011). From factor analyses, it has further emerged that SOP may consist of two discrete components; adaptive striving and self-criticism (McCreary et al., 2004; O'Connor et al., 2010). This may have led to a cancellation effect where SOP, in this adaptive form, may buffer evaluative concerns. Therefore, SPP may represent a stronger cognitive vulnerability to depression than SOP, particularly in the context of achievement stressors.

A third hypothesis is that the stressor is qualitatively different, depending on one's perfectionism style. Exams may for example consist of multiple stressors, where certain aspects of the event resonate more with a specific trait of perfectionism than others. For pupils who put pressure on themselves to do well, the performance during the assessment day may be the most stress-inducing aspect of the exam. Once this stressor has passed, they may be able to start the process of accepting negative outcomes (e.g. the perception

of their performance as sub-standard). For pupils who experience demands to come from others, the main stressor may however not be the exam, but rather the fear of being judged once they have received their grades. Hence, for pupils with high levels of SPP the stressor may still be ongoing until after the results of the exams are known, because of the anticipated negative social consequences.

4.3. **Limitations**

Some limitations should be taken into consideration when interpreting the findings from the present study. Firstly, the alpha values selected for the data analysis were not adjusted to correct for multiple comparisons, which would increase the risk of false positive findings (Type I errors). This decision was weighed up by considering the risk of false negative findings (Type II errors), which in the present study were judged as substantial due to the anticipation of small effect sizes, and a relatively modest sample size. Furthermore, the risks associated with a Type II error could arguably have a more detrimental outcome than if a Type I error was made; e.g. to disregard a small, albeit present effect of perfectionism and not offer pupils support at a period when it would be the most needed, compared to the risk associated with recommending support for selected pupils, who may not have been in more need of support than their peers. The decision to maintain an alpha value of 0.05 was justified by considering that hypotheses were specific, made *a priori*, and had a solid grounding in the literature (Armstrong, 2014).

A second limitation was that measures of perfectionism and rumination were collected only at Time 1. This decision was grounded on theories conceptualising perfectionism and rumination as trait-like, and hence stable over time (e.g. Flett et al., 2016; Abela et al., 2007). It was also made for practical reasons. Specifically, the second screening had to be brief to fit in with school's planned end-of-term activities and to avoid the screenings having an arduous effect on the pupils. It is however possible that certain circumstances, such as exams, are likely to activate more cognitions and behaviours linked to perfectionism or rumination, or to bring them at the forefront of pupils' minds. A lack of Time 2 measures meant that test-retest reliability could not be assessed in the sample. Furthermore, a direct relationship between variables could not be explored at Time 2, limiting our understanding of the discrepancies between the cross-sectional

(Time 1) and longitudinal (Time 2) results of this study. Moreover, the timing of data collection should be considered. At three weeks before exams, pupils' baseline scores of depression were likely to be elevated due to the proximity of the stressor, but it remains unclear when this distress begins to intensify. Despite these limitations, this study is a valuable step towards better understanding the links between perfectionism, rumination and depression in the context of achievement stressors, which evokes questions and presents possible directions for future research.

Finally, we do not wish to portray the relationship between perfectionism, rumination and depression in an overly simplistic model. This study has helped to broaden our understanding of how perfectionism and rumination interact and may affect the trajectory for pupils with high SPP. However, many other cognitive processes and emotions are likely be intertwined, which may have detrimental or buffering effects depending on the context, for example self-esteem and help-seeking. It is therefore important to continue expanding this model, and explore how young people build resilience to these types of stressors, which can guide interventions for those who find exam periods harder.

4.4. **Implications for research**

This study has highlighted some implications for further research. Firstly, it has added to the argument that more longitudinal studies on perfectionism are needed (Morris & Lomax, 2015). In the present study, differences between SOP and SPP in their prospective associations to depression could be found after only five weeks. The use of exams as an achievement stressor was also found to be a helpful paradigm to explore such differences, as the stressor was predictable, universal, and believed to be especially potent for young perfectionists.

Consistent with previous literature, rumination was found to be an important mediator between perfectionism and depressive symptoms in the present study. Future research should continue to explore other possible mediating or moderating variables influencing this relationship. Cognitive factors, such as hope, environmental factors, such as discrepancies between actual and perceived performance, or behavioural indicators of perfectionism could be considered. Hewitt and colleagues (2011) for example suggested

that perfectionism can be observed as self-promotion (i.e. trying to present oneself as perfect), non-display of imperfection (i.e. fear of making mistakes in front of others) or non-disclosure of imperfection (i.e. keeping difficulties a secret from others). A better understanding of how such variables interact, could also improve our knowledge of why perfectionism may place some young people at a greater risk of developing depressive symptoms, but not others.

The present study could furthermore be expanded to clinical samples. Most young people who are seen in mental health services are still enrolled in the education system, hence also have exams. We may presume that for these people, emotional resources are already stretched, which could make exam periods a particularly difficult time. These risks have been highlighted by charities such as Young Minds (2011), but to our knowledge, this has not been well researched in academia.

Finally, we offer some practical considerations for future research. In the present study, a barrier to more schools taking part in the project was the time requirement. Screenings can indeed be time-consuming and impractical. An important area of research will therefore be to tweak and shorten assessment tools, yet ensure reliability is maintained. The use of modern technology, such as mobile phone apps should be considered, as this may help increase response rates and reduce the administrative burden. Apps are also a medium of information gathering that young people nowadays are accustomed to.

4.5. **Clinical implications**

The overall increase in depressive symptoms before the exams has highlighted the importance of strengthening support for pupils during this period, and to offer help proactively. Perfectionism is thought to reduce help-seeking and increase thoughts such as ‘I should cope alone’ (Hewitt et al., 2011). This in turn is thought to raise clinical risk (O’Connor, 2007; Hoff & Muehlenkamp, 2009). It will therefore be particularly important to encourage pupils to seek support when needed, for example by promoting school resources, providing information about helplines and emergency numbers, and by asking parents and school staff to be especially observant of changes in mood or behaviours during this period.

In line with previous recommendations by Flett et al. (2011) and Olson and Kwon (2008), this study also presents evidence emphasising the importance of developing cognitive behavioural interventions which target rumination in clinical work with young perfectionists. Flett and colleagues (2011) further recommend a need for interventions specifically targeting perfectionism as 'the root of their difficulties' (p. 170). For some young people, perfectionism may indeed be a key concept to their psychological formulation and form part of their individual treatment. However, when considering preventative, universal interventions, a focus on broader constructs such as rumination, resilience or assertiveness, which may mediate perfectionism, are likely to be more applicable across all students, and hence also more cost-effective. Psychoeducational groups could for example be offered as part of Personal, Social, Health and Economic (PSHE) educational classes.

Because it appears pupils with high SPP are particularly affected by exams, it will also be important to involve parents and school staff in attempts to reduce environmental pressure. Psychoeducation about the harmful effects of expecting perfection or setting unrealistic expectations should be provided. It may also be helpful to assist parents and pupils alike in having a better understanding of different pathways and options for subsequent education or training (i.e. discussing alternative professions if top grade are unlikely to be achieved).

4.6. **Conclusion**

The present study has added to the picture of rumination as an important mechanism involved in the association between perfectionism and depression. However, when exploring associations prospectively, socially-prescribed perfectionism was found to be a unique and direct predictor of depressive symptoms, in the context of encountering an achievement stressor. These findings may suggest that young people with this type of predisposition may be at a higher risk of experiencing distress following examination periods compared to their peers. More longitudinal studies will be needed to further explore if these changes in mood are maintained, and whether, through repeated exposure to achievement related tasks, the link between SPP and depression may become solidified. A further exploration of other variables which may mediate or moderate the relationship

between perfectionism and depression should also be considered, as they may help explain how this association is aggravated or buffered. Clinical recommendations include promoting awareness about the impact of exam stress on wellbeing, and to proactively offer pupils extra support during exam periods.

References

Abela, J. R. (2002). Depressive mood reactions to failure in the achievement domain: A test of the integration of the hopelessness and self-esteem theories of depression. *Cognitive Therapy and Research*, 26, (4), 531-552.

Abela, J. R., Aydin, C. M., & Auerbach, R. P. (2007). Responses to depression in children: Reconceptualizing the relation among response styles. *Journal of abnormal child psychology*, 35, (6), 913-927.

Abela, J. R., Brozina, K., & Haigh, E. P. (2002). An examination of the response styles theory of depression in third-and seventh-grade children: A short-term longitudinal study. *Journal of Abnormal Child Psychology*, 30, (5), 515-527.

Abela, J. R., & Hankin, B. L. (2009). Cognitive vulnerability to depression in adolescents: A developmental psychopathology perspective. *Handbook of adolescent depression*, 335-376.

Abela, J. R., & Hankin, B. L. (2011). Rumination as a vulnerability factor to depression during the transition from early to middle adolescence: a multiwave longitudinal study. *Journal of abnormal psychology*, 120, (2).

Abela, J. R., Rochon, A., & Vanderbilt, E. (2000). The children's response styles questionnaire. *Unpublished questionnaire*.

Accordino, D. B., Accordino, M. P., & Slaney, R. B. (2000). An investigation of perfectionism, mental health, achievement, and achievement motivation in adolescents. *Psychology in the Schools*, 37, (6), 535-545.

Afshar, H., Roohafza, H., Sadeghi, M., Saadaty, A., Salehi, M., Motamedi, M., ... & Asadollahi, G. (2011). Positive and negative perfectionism and their relationship with anxiety and depression in Iranian school students. *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences*, 16, (1), 79.

Angold, A., and E. J. Costello. "Mood and feelings questionnaire (MFQ)." *Durham: Developmental Epidemiology Program, Duke University* (1987).

Armstrong, R. A. (2014) When to use the Bonferroni correction. Ophthalmic and Physiological Optics, 34, (5).

Ashby, J. S., Dickinson, W. L., Gnilka, P. B., & Noble, C. L. (2011). Hope as a mediator and moderator of multidimensional perfectionism and depression in middle school students. *Journal of Counseling & Development*, 89, (2), 131-139.

Avenevoli, S., Swendsen, J., He, J. P., Burstein, M., & Merikangas, K. R. (2015). Major depression in the National Comorbidity Survey–Adolescent Supplement: prevalence, correlates, and treatment. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54, (1), 37-44.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, 51, (6), 1173.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. University of Pennsylvania Press.

Bibeau, L., & Dupuis, G. (2007). Psychological stress, perfectionism, dissatisfaction and body image: Evaluation of an explanatory model of the issue of suicide. *Canadian Journal of Behavioural Science-Revue Canadienne Des Sciences Du Comportement*, 39, (2), 119-134.

Bieling, P. J., Israeli, A. L. & Antony, M. M. (2004) Is perfectionism good, bad, or both? Examining models of perfectionism construct. *Personality and Individual difference*, 36.

Bieling, P. J., Israeli, A., Smith, J., & Antony, M. M. (2003). Making the grade: The behavioural consequences of perfectionism in the classroom. *Personality and Individual Differences*, 35, (1), 163-178.

Blatt, S. J. (1995). The destructiveness of perfectionism: Implications for the treatment of depression. *American psychologist*, 50, (12), 1003.

Blatt, S. J., & Zuroff, D. C. (1992). Interpersonal relatedness and self-definition: Two prototypes for depression. *Clinical Psychology Review*, 12, (5), 527-562.

Boergers, J., Spirito, A., & Donaldson, D. (1998). Reasons for adolescent suicide attempts: Associations with psychological functioning. *Journal of the American Academy of Child & Adolescent Psychiatry*, 37, (12), 1287-1293.

Burwell & Shirk (2007) Subtypes of rumination in adolescence: Associations between brooding, reflection, depressive symptoms, and coping. *Journal of clinical child and adolescent psychology*, 36, (1), 56-65.

Chartier, M., Stoep, A. V., McCauley, E., Herting, J. R., Tracy, M., & Lymp, J. (2008). Passive versus active parental permission: Implications for the ability of school-based depression screening to reach youth at risk. *Journal of School Health*, 78, (3), 157-164.

Cole, D. A., Jacquez, F. M., & Maschman, T. L. (2001). Social origins of depressive cognitions: A longitudinal study of self-perceived competence in children. *Cognitive Therapy and Research*, 25, (4), 377-395.

Conroy, D. E., Kaye, M. P., & Fifer, A. M. (2007). Cognitive links between fear of failure and perfectionism. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, 25, (4), 237-253.

Costello, E. J., Copeland, W., & Angold, A. (2011). Trends in psychopathology across the adolescent years: what changes when children become adolescents, and when adolescents become adults? *Journal of Child Psychology and Psychiatry*, 52, (10), 1015-1025.

Cyranowski, J. M., Frank, E., Young, E., & Shear, M. K. (2000). Adolescent onset of the gender difference in lifetime rates of major depression: a theoretical model. *Archives of general psychiatry*, 57, (1), 21-27.

Daviss, W. B., Birmaher, B., Melhem, N. A., Axelson, D. A., Michaels, S. M., & Brent, D. A. (2006). Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. *Journal of Child Psychology and Psychiatry*, 47, (9), 927-934.

Denscombe, M. (2000). Social conditions for stress: young people's experience of doing GCSEs. British Educational Research Journal, 26, (3), 359-374.

Dour, H. J., & Theran, S. A. (2011). The interaction between the superhero ideal and maladaptive perfectionism as predictors of unhealthy eating attitudes and body esteem. *Body image*, 8, (1), 93-96.

Education System in the UK (n.d.) Retrieved on 1st October 2017 from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/219167/v01-2012ukes.pdf.

Egan, S. J., Wade, T. D., & Shafran, R. (2011). Perfectionism as a transdiagnostic process: A clinical review. *Clinical psychology review*, 31, (2), 203-212.

Einstein, D. A., Lovibond, P. F. & Gaston, J. E. (2000) Relationship between perfectionism and emotional symptoms in an adolescent sample. Australian Journal of Psychology, 52, (2).

Enns, M. W. & Cox, B. J. (1999) Perfectionism and depressive symptom severity in major depressive disorder. Behavioural Research and therapy, 37, (8), 783-794.

Fergusson, D. M., & Woodward, L. J. (2002). Mental health, educational, and social role outcomes of adolescents with depression. *Archives of general psychiatry*, 59, (3), 225-231.

Flett, G. L., Coulter, L. M., Hewitt, P. L., & Nepon, T. (2011). Perfectionism, rumination, worry, and depressive symptoms in early adolescents. Canadian Journal of School Psychology, 26, (3).

Flett, G. L., Hewitt, P. L., Besser, A., Su, C., Vaillancourt, T., Boucher, D., ... & Gale, O. (2016). The Child–Adolescent Perfectionism Scale: Development, psychometric properties, and associations with stress, distress, and psychiatric symptoms. *Journal of Psychoeducational Assessment, 34*, (7), 634-652.

Flett, G., Hewitt, P. L., Boucher, D. J., Davidson, L. A., & Munro, Y. (2000). The Child-Adolescent Perfectionism Scale: Development, Validation, and Association with Adjustment. Unpublished manuscript.

Flett, G. L., Madorsky, D., Hewitt, P. L., & Heisel, M. J. (2002). Perfectionism cognitions, rumination, and psychological distress. *Journal of Rational-Emotive and Cognitive-Behavior Therapy*, *20*, (1), 33-47.

Flett, G. L., & Hewitt, P. L. (2006). Positive versus negative perfectionism in psychopathology: A comment on Slade and Owens's dual process model. *Behavior modification*, 30, (4), 472-495.

Frost, R. O., Heimberg, R. G., Holt, C. S., Mattia, J. I., & Neubauer, A. L. (1993). A comparison of two measures of perfectionism. *Personality and individual differences*, 14, (1), 119-126.

Frost, R.O., Marten, P., Lahart, C., & Rosenblate, R. (1990) The dimensions of perfectionism. Cognitive therapy and research, 14, (5), pp. 449-468.

Ge, X., Conger, R. D., & Elder, G. H., Jr. (2001). Pubertal transition, stressful life events, and the emergence of gender differences in adolescent depressive symptoms. *Developmental Psychology*, 37, (3), 404-417.

Gladstone, T.R.G., Beardslee, W.R. & O'Connor, E.O. (2011) The prevention of depression. *Psychiatric Clinics of North America*, 34, (1), 35-52.

Greater London Authority (2014) *Ward profiles and atlas.* Retrieved on 5th December 2017 from: https://data.london.gov.uk/dataset/ward-profiles-and-atlas.

Hamachek, D. E. (1978). Psychodynamics of normal and neurotic perfectionism. *Psychology: A Journal of Human Behavior*.

Hammen, C. (2005) Stress and depression. Annual Review of Clinical Psychology,1, 293–319.

Hankin, B. L. (2006). Adolescent depression: description, causes, and interventions. Epilepsy & behavior, 8, (1), 102-114.

Hankin, B. L., & Abela, J. R. (Eds.). (2005). *Development of psychopathology: A vulnerability-stress perspective*. Sage.

Hankin, B. L., & Abramson, L. Y. (2002). Measuring cognitive vulnerability to depression in adolescence: Reliability, validity, and gender differences. Journal of clinical child and adolescent psychology, 31, (4), 491-504.

Hankin, B. L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & Angell, K. E. (1998). Development of depression from preadolescence to young adulthood: emerging gender differences in a 10-year longitudinal study. *Journal of abnormal psychology*, 107, (1), 128.

Hankin, B. L., Fraley, R. C., & Abela, J. R. (2005). Daily depression and cognitions about stress: evidence for a traitlike depressogenic cognitive style and the prediction of depressive symptoms in a prospective daily diary study. *Journal of Personality and Social Psychology*, 88, (4), 673.

Hankin, B. L., Mermelstein, R., & Roesch, L. (2007). Sex differences in adolescent depression: Stress exposure and reactivity models. Child development, 78, (1), 279-295.

Hankin, B. L., & Roesch, L. (2005). Cognitive vulnerabilities to depression and stress: general and specific predictors of psychopathology in youth. *Society for Research in Child Development*.

Hankin, B. L., Young, J. F., Abela, J. R., Smolen, A., Jenness, J. L., Gulley, L. D., ... & Oppenheimer, C. W. (2015). Depression from childhood into late adolescence: Influence of gender, development, genetic susceptibility, and peer stress. *Journal of abnormal psychology*, *124*, (4), 803.

Harris, P. W., Pepper, C. M., & Maack, D. J. (2008). The relationship between maladaptive perfectionism and depressive symptoms: The mediating role of rumination. *Personality and Individual differences*, *44,* (1), 150-160.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.

Hewitt, P. L., Caelian, C. F., Flett, G. L., Sherry, S. B., Collins, L., & Flynn, C. A. (2002). Perfectionism in children: Associations with depression, anxiety, and anger. *Personality and Individual Differences*, *32*(6), 1049-1061.

Hewitt, P. L., Flett, G. L., Turnbull-Donovan, W., & Mikail, S. F. (1991). The multidimensional perfectionism scale: Reliability, validity, and psychometric properties in psychiatric samples. Psychological Assessment, 3, 464–468.

Hewitt, P. L., & Genest, M. (1990). The ideal self: Schematic processing of perfectionistic content in dysphoric university students. *Journal of Personality and Social Psychology*, 59, (4), 802.

Hoff, E. R., & Muehlenkamp, J. J. (2009). Nonsuicidal self-injury in college students: The role of perfectionism and rumination. *Suicide and Life-Threatening Behavior*, 39, (6), 576-587.

Huggins, L., Davis, M. C., Rooney, R., & Kane, R. (2008). Socially prescribed and self-oriented perfectionism as predictors of depressive diagnosis in preadolescents. *Journal of Psychologists and Counsellors in Schools*, 18, (2), 182-194.

IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.

Ingram, R. E. (2003). Origins of cognitive vulnerability to depression. *Cognitive Therapy and Research*, 27, (1), 77-88.

Ingram, R. E. & Luxton, D.D. (2005) Vulnerability-Stress Models. In Hankin, B.L., & Abela, J. R. (Eds) *Vulnerability to psychopathology: Risk across the lifespan*, Chapter 2. Sage.

Ingram, R. E., Miranda, J., & Segal, Z. V. (1998). *Cognitive vulnerability to depression*. Guilford Press.

Ingram, R. E., & Price, J. M. (Eds.). (2010). *Vulnerability to psychopathology: Risk across the lifespan*. Guilford Press.

Keshavan, M. S., Giedd, J., Lau, J. Y. F., Lewis, D. A., & Paus, T. (2014) Changes in the adolescent brain and the pathophysiology of psychotic disorders. *The Lancet Psychiatry*, 1, (7).

Kessler, R. C., Avenevoli, S., & Merikangas, K. R. (2001). Mood disorders in children and adolescents: an epidemiologic perspective. *Biological psychiatry*, 49, (12), 1002-1014.

Kim-Cohen, J., Caspi, A., Moffitt, T. E., Harrington, H., Milne, B. J., & Poulton, R. (2003). Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. *Archives of general psychiatry*, 60, (7), 709-717.

Kirkegaard Thomsen, D. (2006) The association between rumination and negative affect: A review. Cognition and emotion, 20, (8), 1216-1235.

Klein, D. N., Shankman, S. A., Lewinsohn, P. M., & Seeley, J. R. (2009). Subthreshold depressive disorder in adolescents: predictors of escalation to full-syndrome depressive disorders. Journal of the American Academy of Child & Adolescent Psychiatry, 48, (7), 703-710.

Kyriacou, C., & Butcher, B. (1993). Stress in Year 11 school children. Pastoral Care in Education, 11, (3), 19-21.

Lakdawalla, Z., Hankin, B. L., & Mermelstein, R. (2007). Cognitive theories of depression in children and adolescents: A conceptual and quantitative review. *Clinical child and family psychology review*, 10, (1), 1-24.

Lewinsohn, P. M., Joiner Jr, T. E., & Rohde, P. (2001). Evaluation of cognitive diathesis-stress models in predicting major depressive disorder in adolescents. *Journal of abnormal psychology*, 110, (2), 203.

Lewinsohn, P. M., Rohde, P., Klein, D. N., & Seeley, J. R. (1999). Natural course of adolescent major depressive disorder: I. Continuity into young adulthood. *Journal of the American Academy of Child & Adolescent Psychiatry*, 38, (1), 56-63.

Locker, J., & Cropley, M. (2004). Anxiety, Depression and Self-Esteem in Secondary School Children An Investigation into the Impact of Standard Assessment Tests (SATs) and other Im-portant School Examinations. School Psychology International, 25, (3), 333-345.

McCreary, B. T., Joiner, T. E., Schmidt, N. B., & Ialongo, N. S. (2004). The structure and correlates of perfectionism in African American children. *Journal of Clinical Child and Adolescent Psychology*, 33, (2), 313-324.

Meehl, P. E. (1962). Schizotaxia, schizotypy, schizophrenia. *American Psychologist*, 17, 827-838.

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., ... & Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). Journal of the American Academy of Child & Adolescent Psychiatry, 49, (10).

Monroe, S. M., & Hadjiyannakis, H. (2002). The social environment and depression: Focusing on severe life stress. In I. H. Gotlib & C. L. Hammen (Eds.), Handbook of depression (pp. 314-340). New York: Guildford.

Monroe, S. M., Simons, A. D., & Thase, M. E. (1991). Onset of depression and time to treatment entry: roles of life stress. Journal of Consulting and Clinical Psychology, 59, (4), 566.

Morris, L., & Lomax, C. (2014). Assessment, development, and treatment of childhood perfectionism: A systematic review. *Child and Adolescent Mental Health*, 19, (4), 225-234.

Nilsson, K., Sundbom, E., & Hägglöf, B. (2008). A longitudinal study of perfectionism in adolescent onset anorexia nervosa-restricting type. *European Eating Disorders Review*, 16, (5), 386-394.

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of abnormal psychology*, 100, (4), 569.

Nolen-Hoeksema, S., Morrow, J., & Fredrickson, B. L. (1993). Response styles and the duration of episodes of depressed mood. Journal of abnormal psychology, 102, (1), 20.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. Perspectives on psychological science, 3, (5), 400-424.

O'Connor, R. C. (2007). The relations between perfectionism and suicidality: A systematic review. *Suicide and Life-Threatening Behavior*, 37, (6), 698-714.

O'Connor, D. B., O'Connor, R. C., & Marshall, R. (2007). Perfectionism and psychological distress: Evidence of the mediating effects of rumination. *European Journal of Personality*, 21, (4), 429-452.

O'Connor, R. C., Rasmussen, S., & Hawton, K. (2010). Predicting depression, anxiety and self-harm in adolescents: The role of perfectionism and acute life stress. *Behaviour research and therapy*, 48, (1), 52-59.

Office for Standards in Education (2014) Inspection report: Dunraven School HMI 447735. London: OfSTED publications.

Olson & Kwon (2008). Brooding perfectionism: Refining the roles of rumination and perfectionism in the etiology of depression. Cognitive Therapy and Research, 32, (6)

Owen-Yeates, A. (2005). Stress in year 11 students. Pastoral Care in Education, 23(4), 42-51.

Park, R. J., Goodyer, I. M., & Teasdale, J. D. (2004). Effects of induced rumination and distraction on mood and overgeneral autobiographical memory in adolescent major depressive disorder and controls. *Journal of Child Psychology and Psychiatry*, 45, (5), 996-1006.

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry*, 56, (3), 345-365.

PROCESS macro for SPSS. Version 3.0. Retrieved on 12th April 2018 from: http://www.processmacro.org/.

Rice, K. G., & Preusser, K. J. (2002). The adaptive/maladaptive perfectionism scale. *Measurement and evaluation in Counseling and Development*, 34, (4), 210.

Robinson, J. A., Alexander, D. J., & Gradisar, M. S. (2009). Preparing for Year 12 examinations: Predictors of psychological distress and sleep. Australian Journal of Psychology, 61, (2).

Roohafza, H., Talaei, M., Sadeghi, M., Mackie, M., & Sarafzadegan, N. (2010). Association between acute and chronic life events on acute coronary syndrome: a case-control study. *Journal of Cardiovascular Nursing*, 25, (5), 1-7.

Scher, C. D., Ingram, R. E., & Segal, Z. V. (2005). Cognitive reactivity and vulnerability: Empirical evaluation of construct activation and cognitive diatheses in unipolar depression. *Clinical psychology review*, 25, (4), 487-510.

Shafran, R., Cooper, Z., & Fairburn, C. G. (2002). Clinical perfectionism: A cognitive–behavioural analysis. *Behaviour research and therapy*, 40, (7), 773-791.

Sim, H. (2000) Relationship of daily hassles and social support to depression and antisocial behaviour among early adolescents. *Journal of Youth and Adolescence*, 29, (6).

Slaney, R. B., Rice, K. G., Mobley, M., Trippi, J., & Ashby, J. S. (2001). The revised almost perfect scale. *Measurement and evaluation in counseling and development*, 34, (3), 130.

Stikkelbroek, Y., Bodden, D. H. M., Kleinjan, M., Reijnders, M., van Baar, A. L. (2016) Adolescent depression and negative life events, the mediating role of cognitive emotion regulation, *PLoS ONE*, 11, (8).

Stoeber, J., & Rambow, A. (2007). Perfectionism in adolescent school students: Relations with motivation, achievement, and well-being. *Personality and individual differences*, 42, (7), 1379-1389.

Stornelli, D., Flett, G. L., & Hewitt, P. L. (2009) Perfectionism, achievement, and affect in children: A comparison of students from gifted, arts, and regular programs. *Canadian Journal of School Psychology*, 24, (4).

Sund, A. M., Larsson, B., & Wichstrøm, L. (2001). Depressive symptoms among young Norwegian adolescents as measured by the Mood and Feelings Questionnaire (MFQ). *European child & adolescent psychiatry*, 10, (4), 222-229.

Twenge, J. M., & Nolen-Hoeksema, S. (2002). Age, gender, race, socioeconomic status, and birth cohort difference on the children's depression inventory: A meta-analysis. *Journal of abnormal psychology*, 111, (4), 578.

Unger, J. B., Gallaher, P., Palmer, P. H., Baezconde-Garbanati, L., Trinidad, D. R., Cen, S., & Johnson, C. A. (2004). No news is bad news: Characteristics of adolescents who provide neither parental consent nor refusal for participation in school-based survey research. *Evaluation Review*, *28*, (1), 52-63.

Wesselhoeft, R., Sørensen, M. J., Heiervang, E. R., & Bilenberg, N. (2013). Subthreshold depression in children and adolescents–a systematic review. *Journal of Affective Disorders*, 151, (1), 7-22.

Wolfenden, L., Kypri, K., Freund, M., & Hodder, R. (2009). Obtaining active parental consent for school-based research: a guide for researchers. *Australian and New Zealand journal of public health*, 33, (3), 270-275.

Young Minds (2011) *GCSEs: Pressure of exams leaves teens suffering from mental illness.* Published in The Telegraph on 25th August 2011. Retrieved on 2nd May 2018 from: https://www.telegraph.co.uk/education/educationnews/8720513/GCSEs-Pressure-of-exams-leaves-teens-suffering-from-mental-illness.html

Appendix 1. Ethical Approval Letter

Research Ethics Office

Franklin Wilkins Building
5.9 Waterloo Bridge Wing
Waterloo Road
London SE1 9NH
Telephone 020 7848 4020/4070/4077
reo@kcl.ac.uk

KING'S College LONDON

Anna Sofia Musil

21 October 2016

Anna Sofia ,

Study Title: Perfectionism, Rumination, Stress and Depressive Symptoms in Adolescence

Study Reference: Review Reference

I am pleased to inform you that full approval for your project has been granted by the PNM Research Ethics Subcommittee .

For your information, ethical approval has been granted for 3 years from 21 October 2016. If you need approval beyond this point, you will need to apply for an extension at least two weeks before this. You will be required to explain the reasons for the extension. However, you will not need to submit a full re-application unless the protocol has changed.

Ethical approval is required to cover the data-collection phase of the study. This will be until the date specified in this letter. However, you do not need ethical approval to cover subsequent data analysis or publication of the results. For secondary data-analysis, ethical approval is applicable to the data that is sensitive or identifies participants.

Please ensure that you follow the guidelines for good research practice as laid out in UKRIO's Code of Practice for research:
http://www.kcl.ac.uk/innovation/research/support/conduct/cop/index.aspx

Please note you are required to adhere to all research data/records management and storage procedures agreed to as part of your application. This will be expected even after the completion of the study.

If you do not start the project within three months of this letter, please contact the Research Ethics Office.

Please note that you will be required to obtain approval to modify the study. This also encompasses extensions to periods of approval. Please refer to the URL below for further guidance about the process:

http://www.kcl.ac.uk/innovation/research/support/ethics/applications/modifications.aspx

Please would you also note that we may, for the purposes of audit, contact you from time to time to ascertain the status of your research.

If you have any query about any aspect of this ethical approval, please contact the Research Ethics Office:

(http://www.kcl.ac.uk/innovation/research/support/ethics/contact.aspx)

We wish you every success with this work.

Yours sincerely,

James Patterson - Senior Research Ethics Officer

For and on behalf of

Dr Jane Petty, Chair of the PNM Research Ethics Subcommittee

Cc: Eleanor Leigh

Appendix 2. Information Sheet for Parents

INFORMATION SHEET FOR PARENTS AND CARERS

REC Reference Number: HR-15/16-3736

YOU WILL BE GIVEN A COPY OF THIS INFORMATION SHEET

Perfectionism, Rumination, Stress and Depressive Symptoms in Adolescence

Date: 19/10/2017

Dear Parent/Carer,

My name is Sofia Musil and I am training as a Clinical Psychologist at King's College London. I would like to invite your child to participate in this research project which I am undertaking with two highly qualified Clinical Psychologists, Dr Eleanor Leigh (Principal Clinical Psychologist) and Dr Patrick Smith (Consultant Clinical Psychologist and Senior Lecturer). Participation in the project is entirely voluntary and your child should only participate if you and they want to; choosing not to take part will not disadvantage them in any way. Before you decide whether you want your child to take part, it is important that you understand why the research is being done and what your child's participation will involve. Please read the following information carefully and discuss it with others if you wish. Please feel free to contact me (my details are at the bottom) if you would like more information about the project, or if anything is not clear.

What is the purpose of the study?

The study aims to investigate how differences in the way that young people think affect their mood. More specifically, the research project has two main aims:
1) To examine if perfectionism, rumination, and symptoms of depression are connected
2) To investigate if symptoms of depression change over the course of the school year, as academic pressure increases

By exploring these links, we hope this project will help us to better understand why some young people are at a greater risk of experiencing symptoms of depression. This will help us to become better at identifying adolescents who are at risk of becoming depressed, and to improve strategies for preventing and treating depression in young people.

Why has your child been invited to take part?

We are inviting Years 11 pupils in your child's school to take part. We know that many people experience depression for the first time when they are adolescents, so this is a really important period for understanding how it develops. Because of the end of the year exams, pupils in this year group are also under increased academic pressure.

Does my child have to take part?

Participation is completely voluntary. Your child does not have to take part. If you are happy for your child to take part in this research you do not need to do anything further. If you **do not agree** to your child taking part please let us know by Thursday 2nd November 2017 using the attached form or by contacting the research team. If you have any questions please contact the research team using the contact details at the bottom of this sheet.

The project requires children to read and answer a number of questions about their emotions and personality. If you feel your child will be unable to do this because of language or learning difficulties, please let us know. If you are unsure and would like to discuss this or anything else with us please do not hesitate to contact us.

What will happen if my child does take part?

If your child is happy to take part in the study, and you have not informed us that you **do not** wish for your child to take part, your child will be asked to sign a consent form. During your child's form time, we will ask all pupils who have agreed to take part in the study to complete three short questionnaires asking about their mood and thinking. Your child will also be given one short questionnaire in the end of term, which again will ask them about their mood.

At each stage the researchers will be available to answer any questions you or your child may have. The answers that your child provides will be transferred to a database and stored anonymously using a participant number to ensure that your child cannot be personally identified. All data will be anonymised in any reports or publications that result from the project.

Even after starting the project, if you or your child decide that you no longer wish to take part in the project you are free to withdraw your child and all their data from the study at any time until 1st May 2018 without giving any reason. All information relating to them up until that point will be withdrawn from the study and destroyed.

What kind of questions will be asked?

The study will include three questionnaires. The first one will look at perfectionism. Your child will be given a list of sentences such as "I feel that I have to do my best all the time" or "I get

mad with myself when I make a mistake". They will be asked to rate how well each sentence applies to them by circling a number from "1" (False, not at all true of me) to "5" (Very true of me).

The second questionnaire looks at what your child does when they are feeling low or sad. It will ask questions such as "When you feel sad, do you think about how alone you feel" or "When you feel sad, do you remind yourself that this feeling will go away". Your child will be asked to give an answer by circling a number from "1" (Never) to "5" (Always).

The third questionnaire will look at how pupils have been feeling over the past 2 weeks. It will ask questions such as "I felt miserable" or "I didn't enjoy anything at all". Pupils will be asked to give an answer by circling "Not true", "Sometimes true" or "Always true"

What are the possible risks of taking part?

The main disadvantage to taking part in the study is that your child will be donating time out of their form time to take part. The questionnaires are brief and should take no longer than 5-10 minutes to complete.

The questionnaires will ask your child about their mood. Though unlikely, it is possible that thinking about their feelings could cause them slight distress. We will be available in your child's school at the time of filling in the questionnaire, or we can be contacted afterwards to discuss any concerns you or your child may have. You may withdraw your child from the study or they may choose to discontinue at any time.

We will check your child's responses to the questionnaires immediately after they complete them. If they indicate anything that causes concern, we will discuss this with them in the first instance. We may also contact you, their form tutor or another professional if the risk is significant. This is important so that we can take steps to ensure that they are safe and can gain access to support if necessary. Only the small research team will be able to see the young people's responses and these will not be shared with any pupils or staff at the school or anyone else outside the research team.

What are the possible benefits of taking part?

We hope that the project will benefit your child by increasing awareness and recognition of the symptoms of depression, both individually and more generally within the school. If a child is found to report significant symptoms of depression they will be offered support and information for accessing help.

We will continue to be involved with your child's school throughout the school year, and might assist in wellbeing workshops, PSHE lessons and/or hold psychology careers session for any pupils who are interested in Psychology as a career.

Will my child's taking part be kept confidential?

What your child reports on the questionnaires is regarded as strictly confidential. We will collect questionnaires as your child completes them and they will not be shown to anyone else. Anonymity of the material will be protected by using participant numbers and not your child's name. There will be no possibility of your child as an individual being linked with the data at any time.

Data will be only accessed by the researchers named in this information sheet. All questionnaires will be stored in a locked cabinet within King's College London for 7 years after the completion of the project before being destroyed. All digital data will be anonymised and held on password-locked computer files. Anonymised data may be shared with other researchers at King's College London.

Participation is entirely voluntary. If you or your child asks us to withdraw their data at any time before we submit the projects on 1st May 2018 we will remove all traces of it from the records. You will not need to provide a reason for withdrawing your child from the project.

What will happen to the results of the study?

We will produce a final report summarising the main findings, which can be sent to you and will be shared with the school. We also plan to disseminate the research findings through publication and conferences.

Who should I contact for further information?

If you have any questions or require more information about this study, please contact us using the following contact details:

Researcher contact details:

Sofia Musil
Doctorate in Clinical Psychology
Department of Psychology
Institute of Psychiatry, Psychology & Neuroscience
Addiction Sciences Building
4 Windsor Walk
London, SE5 8AF
sofia.musil@kcl.ac.uk

Supervisor contact details:

Dr Eleanor Leigh and Dr Patrick Smith
N&S Child & Adolescent Mood Disorder Service
Michael Rutter Centre
Maudsley Hospital
De Crespigny Park
London
eleanor.leigh@kcl.ac.uk

School Contact Details:

What if I have further questions, or if something goes wrong?

If this study has harmed your child in any way or if you wish to make a complaint about the conduct of the study you can contact King's College London using the details below for further advice and information:

Dr Eleanor Leigh and Dr Patrick Smith
N&S Child and Adolescent Mood Disorder Service
The Michael Rutter Centre for Children and Young People
Maudsley Hospital
De Crespigny Park
London
SE5 8AZ

Appendix 3. Consent Sheet for Parents

Thank you for reading this information sheet and for considering taking part in this research.

FORM TO BE COMPLETED BY PARENTS WHO DO NOT AGREE TO THEIR CHILD TAKING PART IN RESEARCH STUDY

Please read the Information Sheet about the research and complete this form if you DO NOT agree to your child taking part in the research project.

Title of Study: Perfectionism, Rumination, Stress and Depressive Symptoms in Adolescence

King's College Research Ethics Committee Ref: HR-15/16-3736

Thank you for considering whether you agree to your child taking part in this research. If you have any questions arising from the Information Sheet, please contact the researcher before you decide whether or not to let your child join in. You will be given a copy of this Consent Form to keep and refer to at any time.

Please tick

1. **I confirm that I have read and understood the information sheet dated 22/05/2017 for the above study. I have had the opportunity to consider the information and asked questions which have been answered satisfactorily.** ☐

2. **I DO NOT agree to my child taking part in the above study.** ☐

_____

Child's name

_____ _____ _____

Your Name **Date** **Signature**

Appendix 4. Information Sheet for Young People

INFORMATION SHEET FOR YOUNG PEOPLE

REC Reference Number: HR-15/16-3736

YOU WILL BE GIVEN A COPY OF THIS INFORMATION SHEET

Perfectionism, Rumination, Stress and Depressive Symptoms in Adolescence

Dear Pupil,

My name is Sofia Musil and I am training as a Clinical Psychologist at King's College London. I am supervised by two highly qualified Clinical Psychologists, Dr Eleanor Leigh and Dr Patrick Smith. We would like to invite you to take part in this research project. Participation in this project is entirely voluntary and you should only participate if you want to. If you do not want to you will still take part in school as normal.

Please read this letter to understand what the project is about and to help you decide whether you would like to take part. Please feel free to talk to other people if you want help in deciding whether or not to take part. Please also feel free to ask your form tutor or contact me (my details are at the bottom) if there is anything in this information sheet that is not clear, or if you would like more information about the project.

What is the purpose of the study?

The study looks at how differences in the way that young people think affect their mood. More specifically, we are looking at perfectionism (which is when people set high standards for themselves), rumination (which is when people spend a lot of time thinking about how they feel) and mood. We also want to investigate if pupils' mood typically changes over the course of the school year as academic pressure increases.

We hope this project will help us to better understand why some young people feel depressed. This will help us to become better at identifying adolescents who are at risk of becoming depressed, and to improve our prevention and intervention strategies for depression.

Why have you been invited to take part?

We are inviting Years 11 pupils in your school to take part in this study. This is because we know that for many people who go onto develop depression they will have their first difficulties with their mood in their teenage years. These are also important school years, and because of exams at the end of the year, pupils in these year group can feel they are under a lot of academic pressure.

Do I have to take part?

Participation is 100% voluntary. You do not have to take part if you do not want to. If you have any questions you should ask one of the researchers who will be happy to talk to you. You should not agree to take part in this research until you have had all your questions answered. Each time we come to the school we will ask you if you still want to take part and you can say no at any time without any effect. You do not need to give a reason for not wanting to take part.

The research project requires you to read and answer a number of questions about your feelings and experiences. If you think this might be too hard because of language or learning difficulties, please let us know and you do not have to take part. If you are unsure and would like to talk about this or anything else we are always happy to talk to you. Please either find us when we are in your school or contact us using the information at the end of this letter. You might want to ask an adult to help you do this.

What will happen if I take part?

If you decide to take part you will be given this information sheet to keep, and we will ask you to sign a form stating that you are happy to take part. Your parents are also being given information about the study and will be asked to let us know if they do not wish for you to take part in the project.

We will come to visit your school twice. During form time, we will ask everyone who has agreed to take part in the study to complete 3 questionnaires: one is about your mood and two are about your way of thinking. During our second visit you will be asked to fill in another questionnaire about your mood. On both occasions, the researcher will be available in the classroom to answer any questions you may have. If you are not taking part you will continue with normal school activities.

The answers that you give on the questionnaire will be kept private. We will enter your answers into a spreadsheet. Your name will not be included so that no one apart from the research team will know that they are yours.

Even after starting the project, you can decide you do not want to take part anymore. You can decide that you do not want any information about you to be used in the project up until 1st May

2018. All information relating to you will be removed from the study and destroyed. You do not need to give a reason for not wanting to take part.

What kind of questions will be asked?

The study will include three questionnaires. The first one will look at perfectionism. You will be given a list of sentences such as “I feel that I have to do my best all the time” or “I get mad with myself when I make a mistake”. You will be asked to rate how well each sentence applies to you by circling a number from “1” (False, not at all true of me) to “5” (Very true of me).
The second questionnaire looks at what you do when you are feeling low or sad. It will ask questions such as “When you feel sad, do you think about how alone you feel” or “When you feel sad, do you remind yourself that this feeling will go away”. You will be asked to give an answer by circling a number from “1” (Never) to “5” (Always).

The third questionnaire will look at how pupils have been feeling over the past 2 weeks. It will ask questions such as “I felt miserable” or “I didn’t enjoy anything at all”. Pupils will be asked to give an answer by circling “Not true”, “Sometimes true” or “Always true”

What are the possible disadvantages of taking part?

The main disadvantage to taking part in the project is that you will be giving up a small amount of your form time to take part.

The questionnaires ask about experiences that may be difficult for you and about your mood and feelings at the time of the project. Some, but not all, young people might find these difficult to talk about. We will be available at the time and can be contacted afterwards if you are upset or worried about any of the questions we ask. We will also check everyone’s answers on the questionnaires to see if there are any worries you might have and to make sure you are not at risk of anything bad happening. If we noticed something that makes us think you or someone you know may be at risk of harm then we may need to discuss this with your parent or a school staff member. This is important so that we can take steps to make sure that you are safe. We will always try to talk with you first.

What are the possible benefits of taking part?

We hope that you will find the project interesting and that it will help you be more aware of your feelings and how they are influenced by your thoughts and experiences. We also hope that it will help your school by increasing awareness about depression.

If you have difficulties with your mood or feel low or depressed you will be offered support and information for accessing help.

We will also be doing an assembly or a psychology careers session in your school and would love for you to join us if you are interested in learning more about what psychologists do.

Will my participation be kept private?

What you say on the questionnaires will be kept completely private. We will collect the questionnaires as soon as you complete them and will not show them to anyone else. Your name will not be used on the questionnaires. We will use a secret number instead so that we know who you are but no one outside of the research team does.

Only the researchers at the end of this letter will be able to look at the answers to the questionnaires. All questionnaires will be locked in a secure cupboard within King's College London for 7 years after the project is finished before being completely destroyed. Any information stored on the computer will not have your name on it. It will be protected by a password that only the research team will know. Data that has been anonymised so that it cannot be traced to you may be shared with other researchers at King's College London.

If you change your mind about taking part, you are free to stop without giving any reason up to 1st May 2018 and have all information about you taken out of the project.

What will happen to the results of the study?

After your school has taken part in the project we will write a report about what we have found out. We will share this with your school. We also plan to publish the research in science journals to help others learn from our research.

Who should I contact for further information?

You can speak to your form tutor, Head of year or SENCO. If you would like to speak to us, please contact us using the following contact details:

Researcher contact details:

Sofia Musil
Doctorate in Clinical Psychology
Department of Psychology
Institute of Psychiatry, Psychology and Neuroscience
Addiction Sciences Building
4 Windsor Walk
London, SE5 8AF
sofia.musil@kcl.ac.uk

Supervisor contact details:

Dr Eleanor Leigh & Dr Patrick Smith
N&S Child and Adolescent Mood Disorder Service
Michael Rutter Centre
Maudsley Hospital
De Crespigny Park
London
SE5 8AZ
eleanor.leigh@kcl.ac.uk
patrick.smith@kcl.ac.uk

What if I have further questions, or if something goes wrong?

If this study has harmed you in any way or if you wish to make a complaint about the conduct of the study you or your parents can contact King's College London using the details below for further advice and information:

Dr Eleanor Leigh and Dr Patrick Smith
N&S Child and Adolescent Mood Disorder Service
The Michael Rutter Centre for Children and Young People
Maudsley Hospital
De Crespigny Park
London
SE5 8AZ

Thank you for reading this information sheet and for considering taking part in this research.

Appendix 5. Consent Sheet for Young People

Version Number 2 – 30.09.2016

CONSENT FORM FOR PARTICIPANTS IN RESEARCH STUDIES

Please complete this form after you have read the Information sheet and/or listened to an explanation about the research.

Title of Study: Perfectionism, Rumination, Stress and Depressive Symptoms in Adolescence

King's College Research Ethics Committee Ref: HR-15/16-3736

Thank you for considering taking part in this research. The person organising the research must explain the project to you before you agree to take part. If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

	Please tick
I confirm that I understand that by ticking each box I am consenting to this element of the study. I understand that it will be assumed that unticked boxes mean that I DO NOT consent to that part of the study. I understand that by not giving consent for any one element I may be deemed ineligible for the study.	☐

	Please tick
1. **I confirm that I have read and understood the information sheet dated Version number 2-30.09.2016 for the above study. I have had the opportunity to consider the information and asked questions which have been answered satisfactorily.**	☐
2. **I understand that my participation is voluntary and that I am free to stop or withdraw at any time without giving any reason. Furthermore, I understand that I will be able to withdraw my data up to the point of submission in May 2018.**	☐
3. **I consent to the processing of my personal information for the purposes explained to me. I understand that my information will be handled in accordance with the terms of the UK Data Protection Act 1998.**	☐
4. **I understand that my information may be subject to review by responsible individuals from the College for monitoring and audit purposes.**	☐
5. **I understand that confidentiality and anonymity will be maintained and it will not be possible to identify me in any publications.**	☐
6. **I agree that the research team may use my data for future research and understand that any such use of identifiable data would be reviewed and**	☐

Appendix 6. The Mood and Feelings Questionnaire

Child Self-Report

MOOD AND FEELINGS QUESTIONNAIRE: Long Version

This form is about how you might have been feeling or acting **recently**.

For each question, please check (✓) how you have been feeling or acting ***in the past two weeks***.

If a sentence was not true about you, check NOT TRUE.
If a sentence was only sometimes true, check SOMETIMES.
If a sentence was true about you most of the time, check TRUE.

Score the MFQ as follows:
NOT TRUE = 0
SOMETIMES = 1
TRUE = 2

To code, please use a checkmark (✓) for each statement.	**NOT TRUE**	**SOME TIMES**	**TRUE**
1. I felt miserable or unhappy.			
2. I didn't enjoy anything at all.			
3. I was less hungry than usual.			
4. I ate more than usual.			
5. I felt so tired I just sat around and did nothing.			
6. I was moving and walking more slowly than usual.			
7. I was very restless.			
8. I felt I was no good anymore.			
9. I blamed myself for things that weren't my fault.			
10. It was hard for me to make up my mind.			
11. I felt grumpy and cross with my parents.			
12. I felt like talking less than usual.			
13. I was talking more slowly than usual.			
14. I cried a lot.			

15. I thought there was nothing good for me in the future.			
16. I thought that life wasn't worth living.			
17. I thought about death or dying.			
18. I thought my family would be better off without me.			
19. I thought about killing myself.			
20. I didn't want to see my friends.			
21. I found it hard to think properly or concentrate.			
22. I thought bad things would happen to me.			
23. I hated myself.			
24. I felt I was a bad person.			
25. I thought I looked ugly.			
26. I worried about aches and pains.			
27. I felt lonely.			
28. I thought nobody really loved me.			
29. I didn't have any fun in school.			
30. I thought I could never be as good as other kids.			
31. I did everything wrong.			
32. I didn't sleep as well as I usually sleep.			
33. I slept a lot more than usual.			

Appendix 7. The Child and Adolescent Perfectionism Scale

CAPS

This is a chance to find out about yourself. It is not a test. There are no right answers and everyone will have different answers. Be sure that your answers show how you actually are. Please do not talk about your answers with anyone else. We will keep your answers private and not show them to anyone.

When you are ready to begin, please read each sentence below and pick your answer by circling a number from "1" to "5". The five possible answers for each sentence are listed below:

1 = False—Not at all true of me
2 = Mostly False
3 = Neither True Nor False
4 = Mostly True
5 = Very True of me

For example, if you were given the sentence "I like to read comic books," you would circle a "5" if this is very true of you. If you were given the sentence "I like to keep my room neat and tidy," you circle a "1" if this was false and not at all true of you. You are now ready to begin.

Please be sure to answer all of the sentences.

		False				True
1.	I try to be perfect in everything I do	1	2	3	4	5
2.	I want to be the best at everything I do	1	2	3	4	5
3.	My parents don't always expect me to be perfect in everything I do	1	2	3	4	5
4.	I feel that I have to do my best all the time	1	2	3	4	5
5.	There are people in my life who expect me to be perfect	1	2	3	4	5
6.	I always try for the top score on a test	1	2	3	4	5
7.	It really bothers me if I don't do my best all the time	1	2	3	4	5
8.	My family expects me to be perfect	1	2	3	4	5

		True				False
9.	I don't always try to be the best	1	2	3	4	5
10.	People expect more of me than I am able to give	1	2	3	4	5
11.	I get mad with myself when I make a mistake	1	2	3	4	5
12.	Other people think that I have failed if I do not do my very best all the time	1	2	3	4	5
13.	Other people always expect me to be perfect	1	2	3	4	5
14.	I get upset if there is even one mistake in my work	1	2	3	4	5
15.	People around me expect me to be great at everything	1	2	3	4	5
16.	When I do something, it has to be perfect	1	2	3	4	5
17.	My teachers expect my work to be perfect	1	2	3	4	5
18.	I do not have to be the best at everything I do	1	2	3	4	5
19.	I am always expected to do better than others	1	2	3	4	5
20.	Even when I pass, I feel that I have failed if I didn't get one of the highest marks in the class	1	2	3	4	5
21.	I feel that people ask too much of me	1	2	3	4	5
22.	I can't stand to be less than perfect	1	2	3	4	5

Thank you! ☺

Appendix 8. The Children's Response Style Questionnaire

CRSQ

When kids feel sad, they do and think different things. What about you? What do you do and think when you feel sad? For each question, it is very important that you mark what you usually do, not what you think you should do.

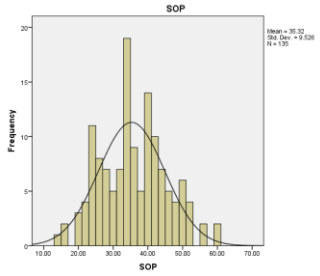
	When you feel sad, do you…	**Almost Never**	**Sometimes**	**Often**	**Always**
1.	Think about how alone you feel	1	2	3	4
2.	Help someone else with something so you don't think about your problem	1	2	3	4
3.	Go away by yourself and think about why you feel this way	1	2	3	4
4.	Watch TV or play video games so you don't think about how sad you are	1	2	3	4
5.	Think "I'm ruining everything"	1	2	3	4
6.	Go to your favourite place and get your mind off your feelings	1	2	3	4
7.	Think about how sad you feel	1	2	3	4
8.	Spend a lot of time on schoolwork	1	2	3	4
9.	Go someplace alone to think about your feelings	1	2	3	4
10.	Do something you enjoy	1	2	3	4
11.	Think about how angry you are with yourself	1	2	3	4
12.	Do something fun with a friend	1	2	3	4
13.	Think about other times when you felt sad	1	2	3	4
14.	Read a book or magazine	1	2	3	4

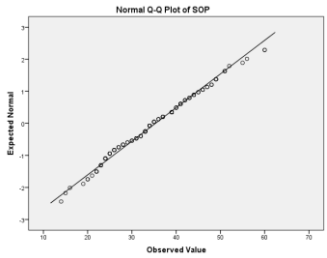
15.	Think about a recent situation wishing it had gone better	1	2	3	4
16.	Ask a friend/parent/teacher to help you solve your problem	1	2	3	4
17.	Think “There must be something wrong with me or I wouldn’t feel this way”	1	2	3	4
18.	Try to find something good in the situation or something you learned	1	2	3	4
19.	Think “I am disappointing my friends/family/teachers”	1	2	3	4
20.	Talk it out with someone who you think can help you feel better	1	2	3	4
21.	Think about all you failures, faults, and mistakes	1	2	3	4
22.	Think of a way to make your problem better	1	2	3	4
23.	Think “Why can’t I handle things better”	1	2	3	4
24.	Remind yourself that this feeling will go away	1	2	3	4
25.	Think about how you don’t feel like doing anything	1	2	3	4

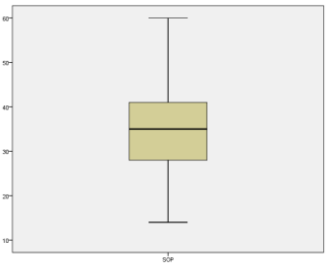
Thank you! ☺

Appendix 9. Visual Representations of data

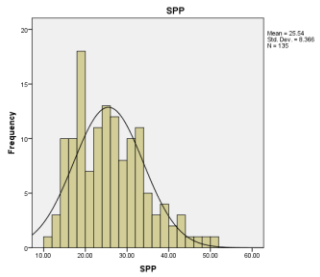
SOP (Self-Oriented Perfectionism)

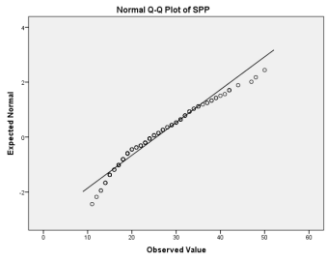


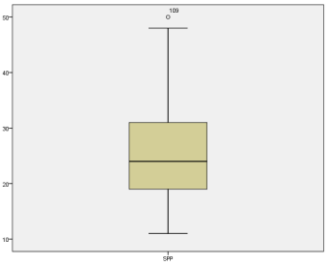




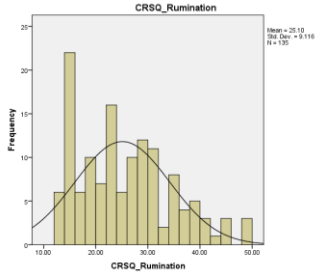
SPP (Socially-Precribed Perfectionism)

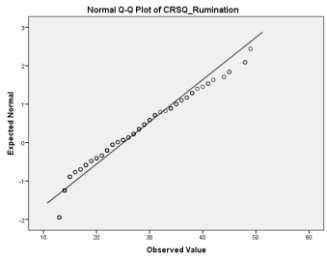


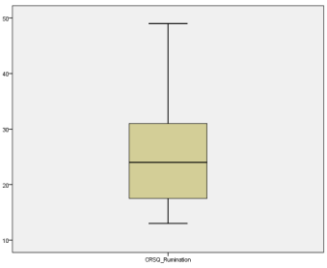




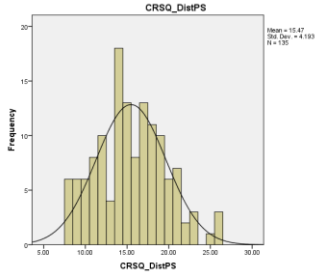
CRSQ-Rumination (Rumination)

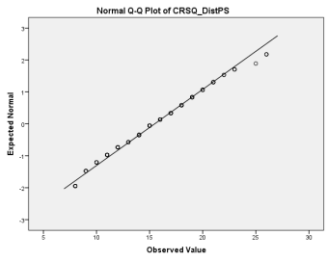


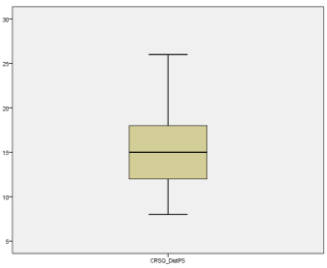




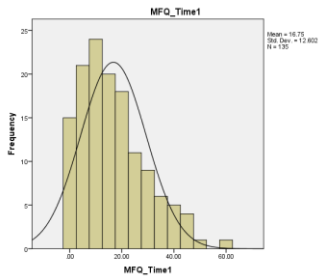
CRSQ-D/PS (Distraction/Problem Solving)

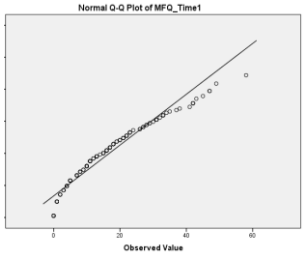


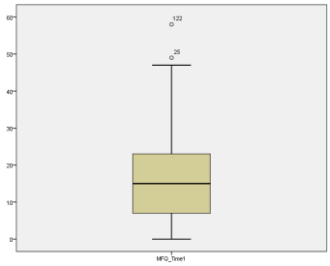




MFQ T1 (Depression at Time 1)

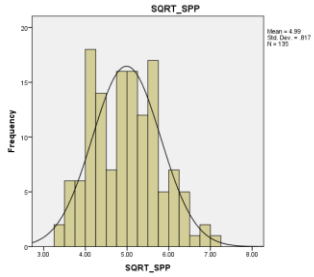


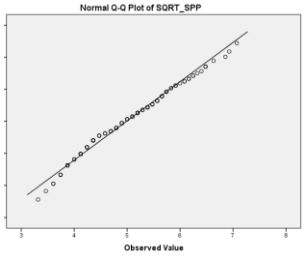


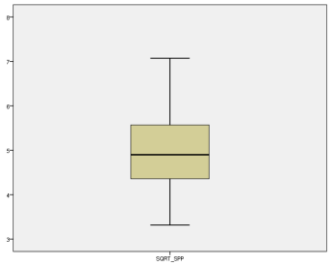


Transformed data

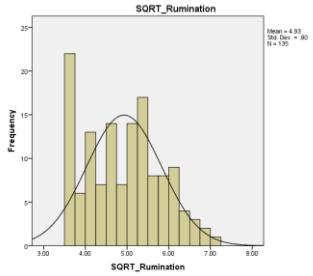
Transformed SPP

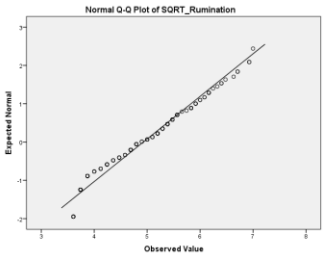




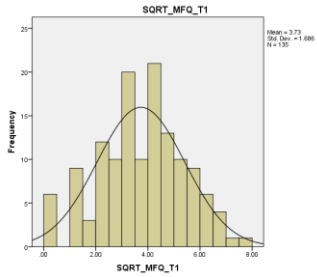


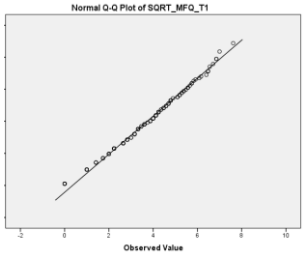
Transformed CRSQ-Rumination

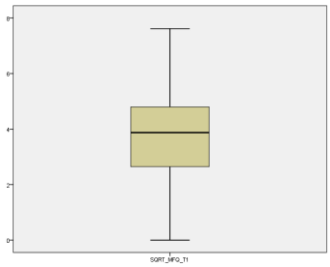




Transformed MFQ T1







SYSTEMATIC REVIEW

Self-report measures of perfectionism developed for children and adolescents: A systematic review

Table of Contents

List of Tables and Figures

Abstract

Objectives

In recent years, perfectionism has been suggested as a cognitive vulnerability linked to mental health difficulties in children and adolescents. The use of valid and reliable screening tools is essential to optimise assessments and evaluate treatment effects. Yet, there is currently no perfectionism instrument which is considered to be of 'gold standard'. The purpose of this review was therefore to critically appraise measures used with children and adolescents.

Methods

Six databases were searched. Articles whose main focus was to evaluate the psychometric properties of a measure in children or adolescent samples were selected according to pre-determined inclusion criteria. The methodological quality of studies was assessed using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist. Psychometric properties were compared against predefined criteria, before findings were summarised using a level of evidence approach.

Results

The search identified 26 articles which met the inclusion criteria. Six instruments had been assessed in the target population. While most studies had explored internal consistency and structural validity, there was a lack of information available regarding test-retest reliability, construct validity and responsiveness of measures.

Conclusion

It will be important for future research to address this gap in the literature, to enable researchers and clinicians to make informed choices about which instrument to use. Tentative recommendations based on the current levels of evidence are presented.

1. Introduction

1.1. Perfectionism in Young People

Perfectionism can be defined as excessive striving towards high standards (Frost, Marten, Lahart & Rosenblate, 1990). Largely recognised as a multidimensional construct (Morris & Lomax, 2014; Hewitt & Flett, 1991; Frost et al., 1990), it has been suggested to encapsulate both adaptive and maladaptive properties (Stoeber & Otto, 2006). Perfectionism has for example been thought to yield a sense of pride and achievement (e.g. Bieling, Israeli, Smith & Antony, 2003), especially among academically gifted children. In stark contrast, there is increasing evidence that these types of traits and cognitions may have detrimental effects on young people's mental health (Morris & Lomax, 2014; Shafran & Mansell, 2001), and links have been found to depression (Huggins, Davis, Rooney & Kane, 2008; Stornelli, Flett & Hewitt, 2009), anxiety (e.g. Hewitt et al., 2002; Stornelli et al., 2009), deliberate self-harm (O'Connor, Rasmussen & Hawton, 2010; Nock & Prinstein, 2005) and eating disorders (e.g. Dour & Theran, 2011). Perfectionism may furthermore increase clinical risk, as it has been found to make young people less likely to seek help in times of distress (e.g. Flett & Hewitt, 2013), and to increase suicidal thoughts and behaviour (Jones, Ramirez, Davies, Canino & Goodwin, 2008).

1.2. Instruments to Measurements of Perfectionism

To identify, prevent and provide treatment when perfectionism becomes of concern, valid, reliable and age-specific measures, such as self-report questionnaires, are essential (Rice & Preusser, 2002). Yet, there is no perfectionism measure which is considered to be of "gold standard". Reflecting the use of different models, several instruments have become available to clinicians and researchers. An overview was shared by Morris and Lomax (2014), who concluded that while having a choice between measures is useful, it also presents difficulties with interpreting and comparing studies. It furthermore makes the selection of instruments difficult, and clinicians or researchers may risk being guided by personal preference or availability of the measure, rather than making decisions based on evidence of psychometric properties (Cheng, Chong & Wong, 1999).

1.3. Rationale of Current Systematic Review

The primary aim of this systematic review was to provide an overview of the validity and reliability of self-report instruments of perfectionism, which had been used with children and adolescents in the field of mental health. Recently, Leone and Wade (2017) became the first authors to publish such a review in English. Some important limitations and differences to the present study were however noted. Firstly, it was not clear if the methodological quality of studies included in their review had been assessed, and if so whether these appraisals were taken into consideration when comparing and summarising findings of psychometric properties. In the present review, a 'level of evidence' approach was instead used (van Tudler, Furlan, Bomberdier & Bouter, 2003). This included systematically rating studies by their methodological rigor, before reported findings were compared. To conclude that the evidence of adequate psychometric properties for an instrument was strong, findings needed to be consistently reported across several studies of 'good' or 'excellent' methodology.

Leone and Wade (2017) also limited their review to only include studies in which the research participants had been children below 15-years of age. Adolescence is believed to be the period in which personality traits, including perfectionist ideals, become more stable (Rice, Leever, Noggle & Lapsley, 2007). For this reason, there is considerable interest in academic and clinical work around perfectionism in this age group, and the appraisal of measures available for the use with adolescents is therefore needed. Moreover, the present review used more than one database to conduct a search for eligible articles.

In conclusion, the present study has aimed to critically appraise the quality of psychometric properties presented for different instruments of perfectionism for young people up to age of 18 years old. By providing a succinct summary of the evidence of each instrument, it is hoped this review will guide clinicians and researchers in deciding which instrument to use.

2. Method

2.1. Search Strategy

The following databases were searched on 1st June 2017: PsychINFO, Embase and Medline (using the OvidSP platform), PubMed, Web of Science (Core collection) and ERIC/ProQuest. Search terms included (Perfection*) AND (child* OR adolescen* OR teen* OR youth OR young) AND (measure* OR assess* OR scale). Wildcards and truncations were used as specified by the different data bases. No publication date or language restrictions were set. A manual search of reference lists was also conducted. The reference management software EndNote X7 (Thomson Reuters, 2016) was used to organise search results. References were exported and processed in Microsoft Excel 2016.

2.2. Selection Criteria

The inclusion and exclusion criteria were as follows:

Inclusion

- The primary purpose of the study was to evaluate the measurement properties (reliability, validity, or responsiveness) of a tool used to appraise perfectionism in a child or adolescent sample
- Participants were 0-18 years old; studies including participants up to 20 years old were included if the mean age was ≤18 (e.g. if the sample was 'high school students')

Exclusion

- The tool assessed had been developed to assess a different construct, where perfectionism featured as a subscale (e.g. The Dysfunctional Attitudes scale; Weissman & Beck, 1978)
- The measure was context specific, for example assessing sport perfectionism only (e.g. Performance Perfectionism Scale for Sport; Hill, Appleton & Mallinson, 2016)
- The primary purpose of the study was not to assess measurement properties *per se*, even if some psychometric measures had been described (for example, studies reporting internal consistency as part of the measurement descriptions)
- The study was not available in English

2.3. Selection Process

All titles and abstracts were screened twice by the first author (SM). Articles which indicated that psychometric properties had been reviewed as a primary aim of the study were selected for full-text review. Selected studies were retrieved through EndNote, OvidSP or Google Scholar. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement and flowchart (Moher et al., 2009) were followed to record results at each stage.

2.4. Overview of Quality Assessment

The quality assessment consisted of four stages, as outlined by the Protocol for Systematic Reviews of Measurement Properties (Terwee, 2011):

Step 1 - Decide which psychometric properties had been assessed by the study

Step 2 - Rate the quality of study using the COSMIN checklist

Step 3 - Compare the property value to a predetermined criterion

Step 4 – Evaluating the instrument using Best Evidence Synthesis

Step 1 – Identifying Psychometric Properties

To decide which psychometric properties had been assessed by the study, the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) taxonomy and definitions were used (Mokkink et al., 2010). The COSMIN taxonomy organises the psychometric properties into three domains; reliability, validity, and responsiveness. Reliability refers to the degree of which the instrument is free from measurement error, whereas validity is the degree to which the instrument is able to capture and measure the intended construct. Responsiveness refers to the validity of a change in score, for example following an intervention. These domains can be further broken into measurement properties, as presented in Table 1. Criterion validity, which can be used to assess if an instrument adequately reflects a ‘gold standard’, was not applicable for this review, as there is no ‘gold standard’ measure for perfectionism.

Table 1. The COSMIN taxonomy, adapted from Mokkink et al. (2010)

Domain	Property	Definition
Reliability	Internal Consistency	The degree of interrelatedness amongst items on an instrument
	Reliability	The proportion of total variance due to 'true' differences between participants
	Measurement Error	Systematic or random error of scores, which can not be attributed to a true change in the construct measured
Validity	Content validity	The degree to which the content of an instrument adequately reflects the construct being measured
	Structural validity	The degree to which scores of an instrument adequately reflects the dimensionality of the construct to be measured
	Hypothesis testing	The degree to which scores on the instrument are consistent with hypotheses (e.g. internal relationships, or relationships to scores on other measures)
	Cross-cultural validity	The degree to which the performance of items on a translated or culturally adapted instrument are an adequate reflection of the performance of the items of the original version of the measure
Responsiveness	Responsiveness	Responsiveness in detecting changes of scores on an instrument

Step 2 - Appraising the Methodological Quality using the COSMIN Checklist

Studies included in the review were appraised using the COSMIN checklist (Mokkink et al., 2010) and 4-point scale (Terwee et al., 2011). The COSMIN checklist has been devised as a modular tool, and the methodological quality of a manuscript is rated separately for each psychometric property addressed (Mokkink et al., 2010). Hence, within a study, quality ratings may vary and be high for some properties but low or absent for others.

The COSMIN 4-point scale consists of nine boxes, one for each psychometric property. Each box includes between 5 and 18 items, which can be rated as 'poor', 'fair', 'good' or 'excellent'. The overall rating for the box is then determined by the lowest score across items (Mokkink et al., 2012). In example, Box A assesses internal consistency. It contains items such as 'was the sample size included in the internal consistency analysis adequate' (item 4) and 'was Cronbach's alpha calculated' (item 9). If the sample size was only moderate and item 4 was rated as 'fair', this would be the overall rating on Box A, even if Cronbach's alpha had been assessed and item 9 had been awarded a rating of 'excellent'. The interpretability and generalisability of the findings were assessed by extracting data about the study characteristics.

To avoid a floor effect in quality ratings, an adaptation to the COSMIN 4-point rating scale was made. Specifically, the rating criteria for one repeated item (which asked if a description of how missing items had been handled had been provided) was relaxed. This was because the item reoccurred in all boxes except from Box D (content validity), but it was only addressed by one of the studies included in the present review. It was therefore agreed that the studies which had been rated as 'fair' on this item, but 'excellent' or 'good' on all other items in the box, could still receive a maximum rating of 'good'.

All studies were rated independently by the first author (SM) and another Trainee Clinical Psychologist (MEJ) who was familiar with the quality assessment tool. Agreement on individual items was 92%, with incongruencies mainly attributed to ambiguity in terminology or rating criteria. Discrepancies between ratings were discussed until the reviewers reached consensus.

Step 3 - Compare the Reported Psychometric Properties to the Pre-Determined Criteria

Study findings were compared against the 'rule of thumb' criteria suggested by Terwee et al. (2007) and Hu and Bentler (1999), to give an indication if the psychometric properties of an instrument could be considered 'acceptable'. The criteria relevant for this

review can be found in Table 2. Narrative summaries have been provided for cross-cultural adaptation of the measures.

Table 2. Quality Criteria for measurement properties, adapted from Terwee 2011

Property	Rating	Criteria
Internal Consistency	+	(Sub)scale unidimensional AND Cronbach's alpha(s) ≥ 0.70
	?	Cronbach's alpha(s) not reported for original structure OR Cronbach's alpha(s) ≥ 0.70 on some but not all subscales
	-	Cronbach's alpha(s) <0.70, despite adequate design and method
Reliability	+	ICC or weighted Kappa ≥ 0.70 OR Pearson's *r* ≥ 0.80 [P,S]
	?	Coefficient not determined OR Doubtful design or method (e.g., time interval not mentioned)
	-	ICC or weighted Kappa <0.70, OR Pearson's *r* <0.80 despite adequate design and method
Structural validity	+	Factors explain at least 50% of the variance OR good or adequate fit by goodness-of-fit criteria for a CFA or EFA [P,GF]
	?	Explained variance not mentioned OR equivocal fit by goodness-of-fit criteria for a CFA or EFA [P,GF]
	-	Factors explain <50% of the variance OR poor fit by goodness-of-fit criteria for a CFA or EFA [P,GF]
Hypothesis testing	+	Specific hypotheses were formulated AND at least 75% of the results are in accordance with these hypotheses
	?	Doubtful design or method (e.g., no hypotheses);
	-	Less than 75% of hypotheses were confirmed, despite adequate design and methods

+ = positive rating, ? = indeterminate rating, - = negative rating; *ICC* = intraclass correlation coefficient;
[P] Park et al.(2013); [S] Schellingerhout et al.(2012); [GF] Good or adequate fit: comparative fit index (CFI) ≥ 0.90, root mean square error of approximation (RMSEA) ≤0.08; Inadequate fit: CFI <0.85, RMSEA >0.10, Indeterminate fit: the values of fit indexes ranged in between the adequate criteria and inadequate criteria (Hu & Bentler, 1999)

Step 4 - Evaluate the Instrument Using Best Evidence Synthesis

The evidence of psychometric properties for each instrument were summarised according to the Best Evidence Synthesis recommendations. These were produced by the Cochrane

Back Review Group (van Tudler et al., 2003) to compare randomized controlled trials, but have been adapted to and used in systematic reviews of self-report measures (e.g. Park, Reilly-Spong & Gross, 2013; Schellingerhout et al., 2012).

A Best Evidence Synthesis (BES) was used, taking into consideration the rigor of studies which assessed the psychometric properties, if the property criteria were met, and the consistency of findings across studies. The level of evidence, rating and criteria have been outlined in Table 3. Because BES is used to compare studies, it is recommended that the study population, setting and version of the instrument is sufficiently similar for findings to be meaningful (Terwee et al., 2011). Therefore, only studies conducted in a community setting and which using the original instrument was included in the evidence synthesis.

Table 3. Level of Evidence for the overall quality of the measurement properties, adapted from van Tudler et al. (2003)

Level	Rating	Criteria
Strong	+ + + or - - -	Consistent findings in at least two studies of good methodological quality, or in one study of excellent methodological quality
Moderate	+ + or - -	Consistent findings in at least two studies of fair methodological quality, or in one study of good methodological quality
Limited	+ or -	Findings reported in one study of fair methodological quality
Conflicting	±	Conflicting findings from studies of comparable methodological quality
Indeterminate	?	Findings from excellent, good or fair studies were not definitely positive or negative
None	na	Findings from excellent, good or fair studies were not available

3. Results

Using the search strategy, a total of 2538 records were identified. Only articles which had indicated that a primary aim of the study was to assess the reliability, validity or interpretability of a perfectionism measure were selected for full-text review. As can be seen in the PRISMA flowchart (Figure 1), some main reasons for exclusion included that the study had used adult samples, that the search terms had identified articles from an unrelated discipline or construct, or that perfectionism was mentioned in the context of specific diagnostic categories. A number of articles (n=296) also assessed perfectionism in children and adolescents, but the main purpose of the article was not to explore psychometric properties. Study which had used cluster analysis only were not included in the review, as these studies did not evaluate measures properties *per se*, but rather provided frameworks for how scores on perfectionism subscales could be categorised.

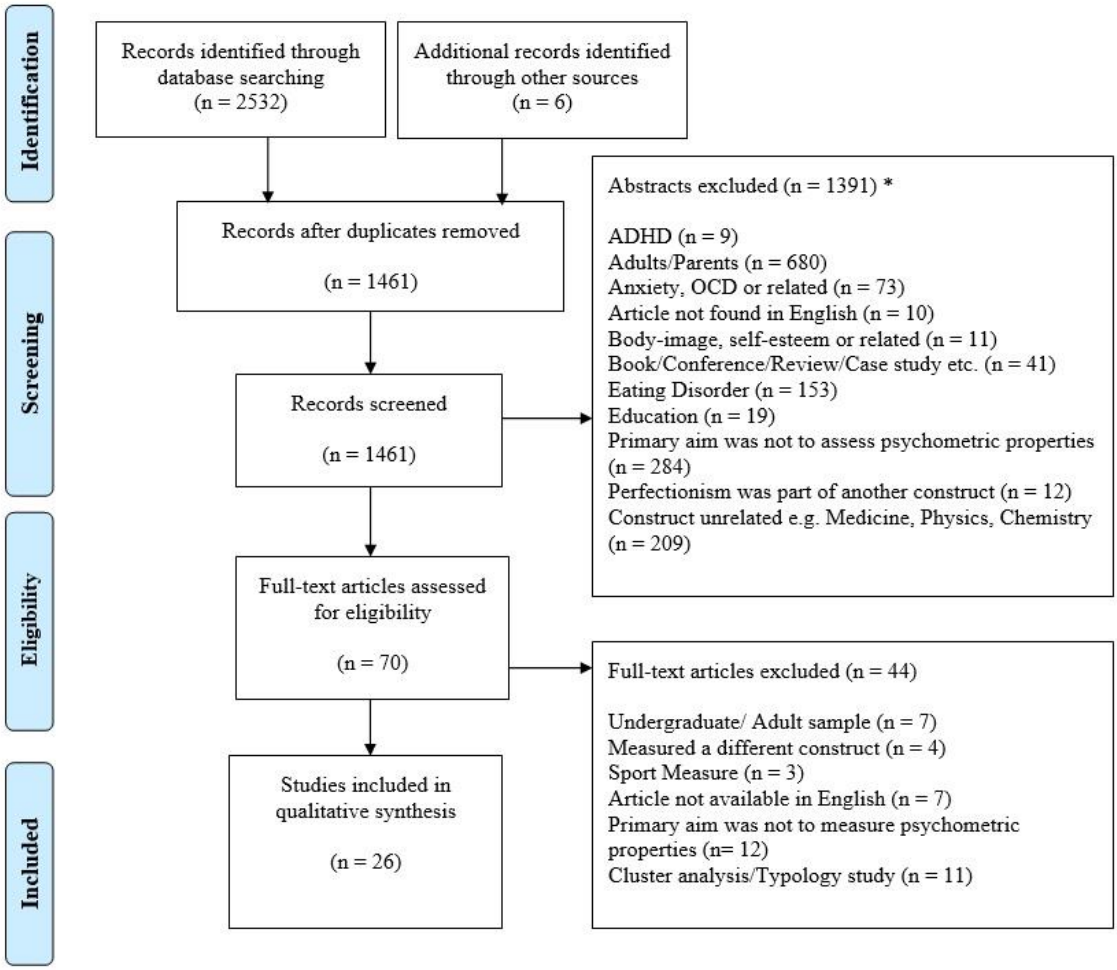


Figure 1 Flowchart of search strategy and selection criteria

*Please note: Some articles were excluded for more than one reason

3.1. Characteristics of included studies

A total of 26 unique articles were identified and included in this review. An overview of the study characteristics can be found in Table 4. All studies had been conducted with community samples. Two studies in addition assessed the use of self-report measures in clinical populations (Castro et al., 2004; Hewitt et al., 2011). Four studies were limited to young people with high academic ability, referred to as 'gifted'. Three studies assessed and compared the psychometric properties of more than one instrument.

Ages of study participants ranged between 8 and 20 years. Ten studies included participants of varied ages (range ≥ 6 years), eight studies had used younger groups of participants (childhood up to pre-adolescence), and seven studies included participants from pre-adolescence up to early adulthood. One study did not report an age-range of their participants.

Twelve of the studies included had been carried out in North America, seven in Europe, two in Australia and five in China. Translated versions of the instruments had been used in 10 of the studies included for review.

Table 4. Characteristics of included studies

Reference	Measure(s)	Population	Age Range	Age Mean (SD)	Sample Size	Female (%)	Country
Baş & Siyez (2010)	CAPS	Community	9-16	11.6 (2.4)	459	58	Turkey
Bento et al. (2014)	CAPS	Community	NR	15.8 (1.5)	971	59	Portugal
Castro et al. (2004)	CAPS (PSPS)	Sample 1 (Community) Sample 2 (Anorexia Nervosa) Subsample (Mix Sample 1 & 2)	11-19 11-19 11-19	14.6 (NR) 15.3 (NR) NR	113 71 68	100 100 100	Spain
Chan (2009)	F-MPS	Community (Gifted)	8-19	12.2 (2.2)	380	40	China
Cheng et al. (1999)	F-MPS	Community	13-18	14.8 (2.0)	947	59	China
Correia et al. (2017)	F-MPS	Sample 1 (Community/Athletes) Sample 2 (Community/Athletes)	12-18 12-18	15.7 (2.5) 15.3 (2.5)	350 206	27 NR	Portugal

Table 4. Characteristics of included studies

Reference	Measure(s)	Population	Age Range	Age Mean (SD)	Sample Size	Female (%)	Country
Flett et al. (2012a)	PSPS-J	Community	Grade 7-8	12.9 (NR)	88	55	Canada
Flett et al. (2012b)	PCI	Sample 1 (Community)	15-19	16.0 (1.0)	250	57	Canada
	CAPS	Sample 2 (Community)	15-19	16.5 (0.8)	105	32	
Flett et al. (2016)	CAPS	Sample 1 (Community)	Grade 3-12	13.3 (3.2)	247	55	Canada
		Sample 2 (Community)	NR	13.5 (4.0)	796	55	
		Sample 3 (Community)	NR	13.0 (0.4)	553	56	
Fong & Yuen (2011)	AMPS	Community	9-13	10.6 (1.2)	599	47	China
Hawkins et al. (2006)	F-MPS	Community (Private School)	Grade 7 & 10	NR	409	100	Australia
Hewitt et al. (2011)	PSPS-J	Sample 1 (Psychiatric)	8-17	13.3 (NR)	244	55	Canada
		Sample 2 (Community)	Highschool	16.3 (NR)	292	48	
		Sample 3 (Cancer survivors)	12-20	15.4 (NR)	65	48	

Table 4. Characteristics of included studies

Reference	Measure(s)	Population	Age Range	Age Mean (SD)	Sample Size	Female (%)	Country
McCreary et al. (2004)	CAPS	Community (African American)	Grade 6	11.8 (0.35)	481	46	USA
O'Connor et al. (2009)	CAPS	Sample 1 (Community)	15-16	15.6 (0.9)	624	52	UK
		Sample 2 (Community)		15.2 (0.7)	514	50	
Parker & Stumpf (1995)	F-MPS	Community (Gifted)	Grade 6	NR	855	37.5	USA
Rice & Preusser (2002)	AMPS	Sample 1 (Community)	9-12	10.1 (0.6)	301	50	USA
		Sample 2 (Community)	9-12	10.4 (0.7)	118	55	
Rice et al. (2004)	AMPS	Community *(Subset from Rice & Preusser, 2002)*	9-11	10.3 (0.6)	113	57	USA
Rice et al. (2007)	AMPS	Community	12-16	13.3 (0.8)	141	57	USA
Sastre-Riba et al. (2016)	APS-R	Community	9-16	12.3 (2.2)	1476	50	Spain
Siegle & Schuler (2000	F-MPS	Community (Gifted)	Grade 6-8	NR	391	58	USA

Table 4. Characteristics of included studies

Reference	Measure(s)	Population	Age Range	Age Mean (SD)	Sample Size	Female (%)	Country
Sironic & Reeve (2015)	F-MPS CAPS APS-R	Community	Grade 9-12	NR	938	62	Australia
Stumpf & Parker (2000)	F-MPS	Community (Gifted)	Grade 6	NR	855	38	USA
Taylor et al. (2017)	F-MPS	Community	12-18	15.6 (1.8)	290	52	UK
Vandiver & Worrell (2002)	APS-R	Community (Gifted)	11-15	13.2 (0.8)	342	52	USA
Wang et al. (2009)	APS-R	Community	14-21	16.6 (1.4)	509	39	China
Yang et al. (2015)	CAPS -C	Community	8-20	13.7 (NR)	933	53	China

NR = Not Reported

3.2. COSMIN ratings

The methodological quality ratings can be found in Table 5. No article selected for review was completely excluded due to lack of poor methodology. Internal consistency was reported by all included articles. Most received the quality rating of 'good'. Lower ratings were typically due to small sample sizes, or a lack of evidence for unidimensionality of the assessed instrument or subscales. Test-retest reliability was assessed in five studies. These were rated to be of 'fair' methodological quality. To receive higher quality ratings, evidence of independent administration, similar test conditions, and evidence that systematic errors in scores had not occurred would have been needed. Measurement error was not assessed by any of the included studies.

Five studies assessed content validity, but only one of the manuscripts was rated to be of 'good' methodological quality. Other studies received a rating of 'poor', mainly due to a lack of involvement from the target population. Structural validity was assessed in 24 studies. The quality of these studies was generally considered 'good'. Lower ratings were predominantly due to small sample sizes or minor flaws in the design or execution of the study. The methodological quality of hypothesis testing varied. Many studies made references to previous findings, but hypotheses were sometimes vague, and it could not be deducted what was expected. Cross-cultural validity was assessed in nine studies. All were rated to be of 'poor' or 'fair' methodological quality, and the translation procedure was typically poorly evidenced. For a higher rating on the COSMIN checklist, more than one forwards and backwards translation of the instrument was required, and the final translation should have been reviewed by a committee (preferably the instrument developers). Responsiveness was not assessed in any of the studies included.

3.3. Characteristics of included Instruments

Six self-report measures of perfectionism were identified. A description of each instrument, including the number of items, subscales, response format and sample items have been presented in Table 6.

Table 5. COSMIN Ratings of Methodological Quality

Measure/ reference	Internal consistency	Reliability	Content validity	Structural validity	Hypothesis testing	Cross-cultural validity
CAPS						
Flett et al. (2016)						
Sample 1	Good		Poor	Good	Fair	
Sample 2	Good			Good	Fair	
Sample 3		Fair				
McCreary et al. (2004)	Good			Good	Good	
O'Connor et al. (2009)						
Sample 1	Good			Good	Good	
Sample 2	Good	Fair		Good	Good	
Sironic & Reeve (2015)	Good			Good	Fair	
CAPS Turkish						
Baş & Siyez (2010)	Good			Good	Fair	Poor
Subsample		Fair				
CAPS Portuguese						
Bento et al. (2014)	Good			Good	Fair	Fair
Subsample		Fair				

Table 5. COSMIN Ratings of Methodological Quality

Measure/ reference	Internal consistency	Reliability	Content validity	Structural validity	Hypothesis testing	Cross-cultural validity
CAPS Spanish						
Castro et al. (2004)						
Sample 1	Poor				Poor	Poor
Sample 2	Poor					
Subsample	Poor	Fair				
CAPS Chinese						
Yang et al. (2015)	Good		Poor	Good	Good	Poor
PSPS-J						
Flett et al. (2012a)	Poor				Good	
Hewitt et al. (2011)						
Sample 1			Poor		Good	
Sample 2	Good			Good	Good	
Sample 3					Poor	
PCI						
Flett et al. (2012b)						
Sample 1	Good			Good	Good	

Table 5. COSMIN Ratings of Methodological Quality

Measure/ reference	Internal consistency	Reliability	Content validity	Structural validity	Hypothesis testing	Cross-cultural validity
Sample 2	Poor					
FMPS						
Hawkins et al. (2006)	Excellent			Excellent	Good	
Parker & Stumpf (1995)	Good			Good	Fair	
Siegle & Schuler (2000)	Good		Poor	Good	Fair	
Sironic & Reeve (2015)	Good			Good	Fair	
Stumpf & Parker (2000)	Poor			Good	Fair	
Taylor et al. (2017)	Good			Good	Good	
FMPS Chinese						
Chan (2009)	Good			Good	Good	Fair
Cheng et al. (1999)	Good			Good	Fair	Poor
FMPS Portuguese						
Correia et al. (2017)						Not Assessed
Sample 1	Fair			Fair	Fair	
Sample 2				Fair	Fair	
APS-R						

Table 5. COSMIN Ratings of Methodological Quality

Measure/ reference	Internal consistency	Reliability	Content validity	Structural validity	Hypothesis testing	Cross-cultural validity
Sironic & Reeve (2015)	Good			Good	Fair	
Vandiver & Worrell (2002)	Good			Good	Fair	
APS-R Spanish						
Sastre-Riba et al. (2016)	Good			Good	Good	Fair
APS-R Chinese						
Wang et al. (2009)	Good			Good	Good	Fair
AMPS						
Rice & Preusser (2002)						
Sample 1	Good		Good	Good	Fair	
Sample 2	Poor			Poor	Good	
Rice et al. (2004)					Good	
Rice et al. (2007)	Good			Good	Good	
AMPS Chinese						
Fong & Yuen (2011)	Good			Good	Fair	Fair

Table 6. Measures of Perfectionism

	Instrument	Items	Response format	Subscales	Sample items
CAPS	The Child-Adolescent Perfectionism Scale (Flett et al. 2000).	22	5-point Likert scale: 1= *False – not at all trues of me* 5 = *Very true of me*	Two subscales; Self-Oriented Perfectionism (SOP), Self-Prescribed Perfectionism (SPP)	• I try to be perfect in everything I do • My family expects me to be perfect • I get upset if there is even one mistake in my work • I feel that people ask too much of me
PSPS-J	Perfectionistic Self-Presentation Scale–Junior Form (Hewitt et al., 2011)	18	5-point Likert scale: 1= *Not at all* 5 = *Extremely*	Three subscales: Perfectionistic Self-Promotion (PSP) Nondisplay of Imperfection (NDP) Non-disclosure of Imperfection (NDC)	• I like trying to look perfect to other people • Mistakes are worse when others see me make them • I should always keep my problems secret
PCI	Perfectionism Cognitions Inventory (Flett et al., 1998)	25	5-point Likert scale to indicate frequency over the past week: 0 = *Not at all* 4= *All of the time*	One Factor	• Why can't I be perfect? • I must be efficient at all times • Maybe I should lower my goals

Table 6. Measures of Perfectionism

	Instrument	Items	Response format	Subscales	Sample items
F-MPS	Frost Multidimensional Perfectionism Scale (Frost et al., 1990)	35	5-point Likert scale: 1= S*trongly disagree* 5 = S*trongly agree*	Six subscales: Personal standards (PS) Concern over Mistakes (CM) Perceived Parental Expectations (PE) Perceived Parental Criticism (PC) Doubting of Actions (DA) Organization (O)	• My parents expect excellence from me • I am a neat person • I expect higher performance in my daily tasks than most people • If I fail in school, I am failing as a person
APS-R	Almost Perfect Scale - Revised (Slaney et al., 2001)	23	7-point scale: 1= S*trongly disagree* 7 = S*trongly agree*	Three subscales: High Standards (HS) Discrepancy (D) Order (O)	• I try to do my best at everything • I am never satisfied with my accomplishments • I think things should be put away in their place
AMPS	The Adaptive/ Maladaptive Perfectionism Scale (Rice & Preusser, 2002)	27	4-point scale: 1= *Really unlike me* 4= *Really like me*	Four subscales: Sensitivity to Mistakes (SM) Contingent Self-Esteem (CSE) Compulsiveness (C) Need for Admiration (NA)	• When I make mistakes I feel so bad I want to hide • Once I do well at something I am pleased • I cannot relax until I have done all my work • I do good work so others think I'm great

3.4. Summary of the Instrument Properties

A descriptive summary of findings for each instrument can be found below. A section for convergence validity reporting correlations between the different measures has also been included. Psychometric properties as reported by each of the studies has been provided in Table 7, followed by a Best Evidence Synthesis in Table 8.

The Child-Adolescent Perfectionism Scale (CAPS)

One of the most widely used measures of perfectionism in children and adolescents has been the CAPS (Flett, Hewitt, Boucher, Davidson, & Munro, 2000). It was intended to measure perfectionistic traits, either in form of the tendency to place high expectations on oneself (self-oriented perfectionism; SOP), or as held beliefs that perfection is demanded from oneself by others (socially-prescribed perfectionism; SPP). The CAPS was adapted from an adult version of the measure, which also included other-oriented perfectionism (placing high expectations on others). This subscale was however not included in CAPS due to uncertainty about when, developmentally, one would expect these traits to manifest in young people (Flett et al., 2016).

In total, eight studies assessing the psychometric properties of CAPS were identified. Four studies assessed the original (English) version of CAPS. One study (McCreary Joiner, Schmidt, & Ialongo, 2004) adapted the response options to a 4-point scale. Three studies assessed translated versions of the measure (Baş & Siyez, 2010; Bento et al., 2014; Castro et al., 2004), and one study (Yang, Hong, Tao, & Zhu, 2015) assessed an extended version of CAPS which had been translated and culturally adapted for a Chinese population.

Reliability

Adequate internal consistencies on both the SOP and SPP subscales were reported by three studies with good methodological quality (Flett et al., 2016; Sironic & Reeve, 2015; Bento et al., 2014). Alpha values for both subscales ranged between 0.81 and 0.87. Meanwhile, two studies found adequate internal consistency for SPP (α >0.82) but not SOP (α =0.55 and 0.64) (McCreary et al., 2004; Baş & Siyez, 2010). Higher internal consistency was indicated for adolescent participants compared to children (Baş and Siyez, 2010). Castro et al. (2004) compared scores on the CAPS between female adolescents with or without anorexia nervosa. Due to small sample sizes the methodological quality of this study was rated as poor; but findings indicated that the

measure had good internal consistency also among their clinical sample ($\alpha = 0.92$ for both SOP and SPP).

Three studies assessed test-retest reliability of the original CAPS, but none of the studies met the property criteria indicated for acceptable reliability. Intraclass correlation coefficients (ICC) of 0.59 (SOP) and 0.69 (SPP) were reported by Bento et al. (2014), following a test interval of five weeks. Test-retest correlations for SOP and SPP were reported for a 2-week period ($r = 0.63$ and 0.73 respectively) by Baş & Siyez (2010), and over one year ($r = 0.65$ and 0.59 respectively) by Flett et al. (2016). A difference in test-retest reliability between age groups was suggested, as participants who first completed the measure when they were in Year 5 reported higher reliability ($r = 0.80$ for SOP and r=0.70 for SPP) than did younger pupils (Flett et al., 2016).

Validity

Content validity was assessed by Flett and colleagues (2016) during the development of the measure, and by Yang et al. (2015) as adaptations to the Chinese-CAPS were made. Neither of the studies involved the target population.

Seven studies assessed the structural validity of the CAPS. The developers (Flett et al. 2016) used exploratory factor analysis (EFA) and reported the two factors accounted for 41% of the variance. Using confirmatory factor analysis (CFA), McCreary et al. (2004) and O'Connor, Dixon and Rasmussen (2009) instead found a three-factor structure to be a better fit, with the SOP subscale divided into SOP-striving and SOP-critical. This model was also supported by Baş & Siyez (2010) using principal component analysis (PCA). These authors have further suggested that the CAPS should be shortened to address cross-loadings. Recommendations of which items to exclude have varied across the studies, but consistently the negatively worded items (items 3,9 and 18) have been found to be particularly problematic. Flett and colleagues (2016) have however argued that these are important to keep for assessment purposes.

When translating the CAPS into Chinese, Yang et al. (2015) used a 2 x 2 model of perfectionism (positive/negative x SOP/SPP), as previously proposed in the adult literature (Chang, 2006). New items were generated to assess positive SPP, and CFA was used to review both the traditional 2-factor model (with the original 22 items), and their new 4-factor model. Only the latter model was found to be a good fit.

Several studies assessed the construct validity of the CAPS by testing for hypothesised associations between perfectionism and related constructs. As predicted, CAPS scores were positively correlated with symptoms of depression (r = 0.19 to 0.60; McCreary et al., 2004; Sironic & Reeve, 2015; Baş & Siyez, 2010) and symptoms of anxiety (r = 0.21 to 0.48; McCreary et al., 2004; Sironic and Reeve, 2015). No statistically significant correlations could be found between the CAPS and symptoms of conduct disorder (McCreary et al., 2004). Yang et al. (2015) reported a positive association between scores on the Chinese CAPS (adapted) and Academic achievement and performance (r = 0.28 to 0.49) and Learning anxiety (r = 0.19), but a negative association with Learning stress (r = -0.27).

Other considerations

Yang et al. (2015) pointed out the importance of considering cultural variation in perfectionism, in particular how expectations from others, in some Asian cultures, may be seen as positive and motivating. This, they argue, applies not only when research is conducted in Asia, but also when interpreting scores from minority groups in Western countries.

The Perfectionistic Self-Presentation Scale–Junior Form (PSPS-J)

Flett, Hewitt and colleagues suggested the perfectionism construct not only encompasses traits, but also behaviours and thought processes. The PSPS-J (Hewitt et al., 2011) was developed to assess interpersonal expression of perfectionism. Three subscales were proposed, including Perfectionistic Self-Promotion (e.g. appearing perfect to others), Non-Display of Imperfection (e.g. fear of making mistakes in front of others), and Nondisclosure of Imperfection (e.g. avoiding telling others about difficulties or mistakes made).

Only two studies were found to have assessed the psychometric properties of the PSPS-J in children or adolescents; both produced by the instrument authors (Hewitt et al., 2011; Flett et al., 2012a). The adult version of the measure was also translated for the use with Spanish adolescents (Castro et al., 2004).

Reliability

Internal consistencies varied across the subscales and were reported as high for Perfectionistic Self-Promotion ($\alpha > 0.89$), adequate for Non-display of Perfectionism ($\alpha = 0.70$ to 0.82), and indeterminate for the Nondisclosure of Imperfection subscale ($\alpha = 0.60$ to 0.72). Test-retest reliability was only assessed by Castro et al. (2004), using the adult version of the measure. An adequate correlation in total PSPS scores was found after one week ($r = 0.83$) in a small subgroup of female participants.

Validity

In the development of the PSPS-J, a pool of items was administered to a clinical sample of young people, and factor analysis was used to decide which items to include in the final version of the instrument (Hewitt et al., 2011). The target population was however not directly consulted about the relevance or quality of the questions.

To assess structural validity of the PSPS-J, Hewitt and colleague (2011) used a second group of participants. Confirmatory factor analysis was used, and the goodness-of-fit indices suggested the proposed 3-factor structure was the best fit for the data.

Consistent with expectations for construct validity, scores on the PSPS-J correlated with scores of measures assessing symptoms of depression ($r = 0.27$ to 0.37; Hewitt et al., 2011), dysfunctional attitudes relating to depression ($r = 0.49$ to 0.66; Flett et al., 2012 a) and social anxiety ($r = 0.47$ to 0.56; Flett et al., 2012 a). Associations were also found between the PSPS-J and several subscales of the Youth Psychopathy Inventory (YPI; Andershed et al., 2002), with the highest correlation reported between Perfectionistic Self-Promotion and disarming charm ($r = 0.42$; Hewitt et al., 2011). As predicted, the strengths of correlation between each of the measures and the three subscales of the PSPS-J varied (Hewitt et al., 2011).

Other considerations

Castro et al. (2004) suggested the PSPS might be of particular relevance to adolescents with Anorexia Nervosa, as focus on appearance forms part of the clinical presentation. In a sample of young women with eating difficulties (Sample 2) they found correlations between the PSPS (adult measure) and abnormal eating behaviours and attitudes ($r = 0.38$; $p < 0.01$).

The Perfectionism Cognitions Inventory (PCI)

Flett and colleagues (Flett, Hewitt, Blankstein & Gray, 1998) developed a third measure of perfectionism, which was intended to capture the frequency of automatic thoughts with perfectionistic content, for example relating to performance or standards. The PCI can therefore be seen as a state measure, as cognitions may be more or less prevalent depending on the situational context (Flett et al., 2012). In contrast to the CAPS and PSPS-J, no adaptation to the PCI was made for the use of younger samples. Only one study assessing the psychometric properties of the PCI was found (Flett et al., 2012).

Reliability

Flett and colleagues (2012) reported high internal consistency of the measure ($\alpha = 0.91$). Test-retest reliability of the PCI was not assessed.

Validity

The content validity of the measure was not assessed in the target population. The factor structure was assessed using principal component analysis. As predicted, only one factor was indicated for the measure.

Higher scores on the PCI were associated with symptoms depression ($r = 0.43$), a higher frequency of negative automatic thoughts ($r = 0.46$), and with scores on the self-criticism ($r = 0.38$) and dependency ($r = 0.39$) subscales of the Adolescent Depressive Experiences Questionnaire (DEQ-A; Fichman et al., 1994). No significant differences in mean scores of the PCI were found between male and female participants. Together, these findings indicated evidence of construct validity of the measure.

The Frost Multidimensional Perfectionism Scale (FMPS)

Another well-established measure is the FMPS (Frost et al., 1990). Based on Hamachek's model of perfectionism, the authors believed that different perfectionist ideals would emerge from an early age and solidify as traits through the reinforcement from the child's environment, for example the perception that high achievements are essential to gain love and approval from care givers. The instrument was first developed and validated with a group of female undergraduate students (Frost et al., 1990), but has also been used with adolescent samples.

Nine studies which assessed the psychometric properties of the FMPS were found. Six of the studies used the original FMPS (small modifications were made to adjust the verb tense from past to present). Three studies used FMPS translated into Chinese (Chan, 2009; Cheng et al. 1999) or Portuguese (Correia, Rosado & Serpa, 2017).

Reliability

The internal consistency was assessed by all studies included in the review, but only three manuscripts (Parker & Stumpf, 1995; Sironic and Reeve, 2015; Taylor, Couper & Butler, 2017) reported alpha values for the original six factors proposed by the measure developers. The evidence of adequate internal consistency was found to be inconclusive, as across the studies Cronbach's alpha was reported as higher than 0.7 on most, but not all subscales. Doubts about actions consistently failed to reach the desired value (α =0.59 to 0.68). Test-retest reliability was not assessed.

Validity

Content validity was assessed in one study (Seigle and Schuler, 2000), but the target population was not involved, and the methodological quality was therefore rated as poor.

The structural validity of the FMPS has been debated. Three studies (Parker & Stumpf, 1995; Stumpf and Parker, 2000; Chan, 2009) reported evidence of an acceptable fit of the six-factors proposed by Frost and colleagues (1990). After excluding two cross-loading items from their analysis (item 16 and 18), Hawkins, Watt and Sinclair (2006) also reported an adequate goodness-of-fit index for the original structure. The authors did however argue that stability of the factors could be increased by merging parental expectations and parental criticism, and concerns about mistakes and doubts about actions. This four-factor solution was supported by Sironic and Reeve (2015). Other proposed solutions have included a one factor (Taylor et al., 2017), five factor (Cheng et al.1999) or eight-factor structure (Siegle & Schuler, 2000). Stumpf and Parker (2000) suggested the subscales could be grouped into healthy and unhealthy forms of perfectionism.

Assessing construct validity, FMPS total scores have been positively correlated with fear of failure (r = 0.52; Correia et al., 2017), emotional problems (r = 0.39; Taylor et al. 2017) and negatively correlated with self-esteem (r = -0.27; Cheng et al., 1999). Correlation

coefficients varied and ranged between r = 0.12 and 0.48 for associations between the different subscales of perfectionism and depression, anxiety and stress, except for the organisation subscale which did not correlate with any of the comparison measures (Sironic and Reeve, 2015).

Other considerations

Cheng et al. (1999) cautioned that parenting style vary between cultures. As Frost's model places emphasis on the parent-child interaction in the development of perfectionism, cultural context may be particularly important when evaluating the use of the FMPS.

The Almost Perfect Scale revised (APS-R)

The APS-R was developed by Slaney and Ashby (1996), after the authors had noted that the multidimensional scales of perfectionism as proposed by Frost et al. (1990) and Flett and Hewitt (Hewitt et al., 1991) predominantly focused on the negative aspect of perfectionism. In contrast, they had found that many people held positive beliefs about perfectionism, uncovered through interviews with young adults (Slaney and Ashby, 1996). The APS-R was therefore developed to measure both adaptive and maladaptive properties of perfectionism (Slaney, Rice, Mobley, Trippi & Ashby, 2001). The measure consists of three subscales; High Standards (striving for high achievement), Discrepancy (reactions when failing to meet standards) and Order (neatness and organisation). The ASP-R was first validated in a sample of undergraduate students (Slaney et al., 2001).

Four studies which assessed the psychometric properties of the APS-R amongst an adolescent population were found in this review. Two manuscripts evaluated the original measure (Vandiver & Worrell, 2002; Sironic & Reeve, 2015), while in two studies the measure had been translated into Spanish (Sastre-Riba, Pérez-Albéniz & Fonseca-Pedrero, 2016) or Chinese (Wang, Yuen & Slaney, 2009).

Reliability

Adequate internal consistency was reported across the three subscales, with alpha values ranging between α = 0.79 and 0.92 (Vandiver & Worrell, 2002; Sironic & Reeve, 2015). Sastre-Riba et al. (2016) also reported adequate internal consistency of the measure when

administered to a large group of Spanish adolescents, when using McDonald's Omega as the coefficient. Test-retest reliability of the APS-R was not assessed.

Validity

Content validity of the APS-R was not evaluated in any of the studies included for review. Structural validity was assessed using factor analysis (EFA and CFA). A three-factor solution was consistently found, but goodness-of-fit criteria was equivocal or could not be met, and in only one study could the proposed solution account for 50% of the variance (Wang et al., 2009).

Sironic and Reeve (2015) found scores on the Discrepancy subscale to correlate with symptoms of depression (r = 0.56/ r = 0.54), anxiety (r =0.49/ r = 0.42), and stress (r = 0.55/ r = 0.49) amongst male and female participants respectively. Weaker correlations were found between High Standards and stress (r =0.20/ r = 0.16) in both gender groups, and between High Standards and anxiety reported by boys (r = 0.18). Wang et al. (2009) similarly found positive correlations between Discrepancy and depression (r = 0.49), and Discrepancy and loneliness (r = 0.37), and negative correlation between Discrepancy and life satisfaction (r = -0.29). In contrast, the Order subscale correlated negatively with depression (r = -0.2) but positively with life satisfaction (r = 0.23). Vandiver and Worrell (2002) administered the Measure of Perceived Life Chances (MPLC; Jessor, Donovan & Costa, 1990), in which participants were asked about their future prospects (e.g., "What are the chances that you will have a job that pays well?"). MPLC scores correlated positively with High standards and Order (r = 0.51 and r = 0.40 respectively) but negatively with Discrepancy (r = -0.25). These findings are consistent with the hypothesis that discrepancy represents a more maladaptive form of perfectionism and indicate evidence of construct validity.

The Adaptive/Maladaptive Perfectionism Scale (AMPS)

AMPS was the first measure of perfectionism to be developed specifically for the use of children (Rice & Preusser, 2002). Similar to the ASP-R, the AMPS is based on a conceptualisation of perfectionism as healthy and unhealthy, as proposed by Adler (1964). The authors suggested these two subtypes could be distinguished by the person's approach to goal-setting (a striving towards attainable or unrealistic targets) and rigidity

of self-evaluation (for example having flexibility to accept some mistakes instead of harsh self-scrutiny regardless of outcomes). Items from various adult instruments were reviewed and adapted for the use by a younger population. Four subcategories were proposed; Sensitivity to Mistakes, Contingent Self-esteem, Compulsiveness and Need for Admiration.

Four studies were identified for this review; two studies assessed validity and reliability of the original measure (Rice & Preusser, 2002; Rice et al., 2007) and one study assessed the AMPS translated into Chinese (Fong & Yuen, 2011). A follow-up study by Rice, Kubal and Preusser (2004) also evaluated the construct validity of the AMPS.

Reliability

Rice and Preusser (2002) reported adequate internal consistency on all four subscales (alpha-values ranged between 0.73 and 0.91). These findings could however not be replicated in a Chinese sample; Fong and Yuen (2011) instead found that $\alpha > 0.70$ only could be reported for the Need for Admiration subscale. Test-retest reliability was not assessed.

Validity

The development study by Rice and Preusser (2002) was the first and only manuscript to receive a methodological quality rating of 'good' for their evaluation of content validity. A pool of 90 questionnaire items were administered to a group of fourth- and fifth-graders, who had the opportunity to comment on the questions and response format. After changes were made, the measure was piloted with another group of pupils. Evidence of content validity was therefore indicated.

Structural validity was first assessed using PCA. A four-factor solution was initially proposed. When the study was replicated with an older sample, the authors instead found a three-factor solution as items on the Contingent Self-Esteem subscale loaded onto the Concerns about Mistakes subscale, and the authors proposed self-criticism may be an underlying factor (Rice et al., 2007). Using CFA, Fong and Yuen (2011) did not find that the goodness-of-fit criteria could be met when all questionnaire items were included in the analysis. Instead they proposed a reduction to 23 items. Three out of the four items

deleted belonged to the Contingent Self-Esteem Subscale, and all had been negatively worded.

To assess construct validity, Rice and colleagues (2004) asked a subsample of pupils (n=113) from their first study to complete the Piers-Harris Self-Concept Scale (PHSCS; Piers, 1969). The PHSCS consists of six subscales; behaviour, intellectual and school status, physical appearance and attributes, anxiety, popularity, and happiness and satisfaction. Sensitivity to Mistakes correlated negatively with most PHSCS subscales (r = -0.23 to -0.63), indicating this dimension of the AMPS was associated with a lower self-concept. Positive correlations were found between most aspects of the PHSCS and the Contingent Self-Esteem subscale (r=0.23 to 0.55), suggesting children who endorsed this subscale of the AMPS were likely to hold higher beliefs about themselves. The Compulsiveness subscale correlated with higher levels of anxiety (r = -0.38/-0.48) and less happiness and satisfaction (r = -0.26/-0.38) amongst boys and girls respectively, and with concerns about physical appearance and attributes (r = -0.35) reported by girls. For boys, Need for Admiration correlated with anxiety (r = 0.27), whereas for girls this subscale of the AMPS was also associated with more concerns about physical appearance and attributes (r = 0.35), and less happiness and satisfaction (r = -0.41). In an older sample of children, no statistically significant correlations between AMPS subscales and symptoms of depression could be found (Rice et al., 2007).

Convergent validity

Five of the studies included for review had used more than one perfectionism measure, reported correlations between the instruments (Hewitt et al., 2011; Flett et al., 2012a; Flett et al., 2012b; Baş & Siyez, 2010; Sironic & Reeve, 2015). The CAPS was used in all studies. Relationships between subscales of the instruments tended to be reported separately, rather than total scores of the instrument. In most cases statistically significant associations were found. The strongest correlations from each study have been summarised below.

CAPS and PSPS-J

CAPS Self-Oriented Perfectionism (SOP) and Self-Prescribed Perfectionism (SPP) both correlated most strongly with the Perfectionistic Self-Promotion subscale of the PSPS-J. Correlation coefficients were reported as r = 0.54 and 0.52 (Hewitt et al., 2011) and r =

0.41 and 0.36 (Flett et al., 2012a) for SOP and SPP respectively. Flett et al. (2012) found a stronger relationship between SOP and Non-display of imperfection (r =0.41) than between SOP and the Non-disclosure of imperfection subscale. The opposite was found for SPP, which correlated more strongly with the Non-disclosure of imperfection subscale (r =0.39). A similar pattern was observed by Hewitt et al. (2011).

CAPS and PCI

Flett et al. (2012b) reported significant correlations between the PCI and SOP (r =0.61), and between PCI and SPP (r =0.50).

CAPS and FMPS

Baş & Siyez (2010) found that most subscales of the FMPS correlated, to varying degrees, with both SPP and SOP (r =0.16 to 0.52). The strongest associations were found between SOP and Personal Standards (r =0.42), and SPP and Parental Expectations (r =0.37). Sironic and Reeve (2015) similarly found SOP to correlated with Personal Standards (r =0.75) and Concerns about Mistakes (r =0.65), whereas SPP instead were more strongly linked to Parental Expectations (r =0.71) and Parental Criticism (r =0.64).

CAPS and APS-R

Sironic & Reeve (2015) found significant associations between SOP and ASP-R High standards (r =0.73), and SPP and ASP-R Discrepancy (r =0.49). ASP-R Order was moderately correlated with SOP (r =0.42), but only weakly with SPP (r =0.12).

F-MPS and APS-R

Strong correlations emerged between the FMPS Personal Standards and ASP-R High Standards (r =0.82), and between FMPS Organisation and ASP-R Order (r =0.87) as reported by Sironic and Reeve (2015). Associations were also found between the FMPS Concerns about Mistakes and Doubts about Actions subscales of the FMPS, and ASP-R Discrepency (r =0.70).

Table 7. Psychometric Properties

Measure/ reference	Internal consistency	Reliability	Structural validity
CAPS			
Flett et al. (2016)			
Sample 1	SOP 0.85 / SPP 0.81		EFA Original 2-factors: 40.9 % variance
Sample 2	SOP 0.81 / SPP 0.84		EFA Proposed 3-factors: 45.5 % variance
Sample 3	-	SOP 0.65 / SPP 0.59 (1 year)	
McCreary et al. (2004)	SOP 0.55 / SPP 0.83	NR: 3-factor only	EFA Original 2-factors: CFI=0.77, RMSEA=0.075 Proposed 3-factors (15-items): CFI=0.88, RMSEA=0.045
O'Connor et al. (2009)			
Sample 1	NR: 3-factor only		CFA Original 2-factors: CFI=0.83, RMSEA=0.08 Proposed 3-factors (14-items): CFI= 0.95 RMSEA=0.057
Sample 2	NR: 3-factor only	NR: 3-factor only	
Sironic & Reeve (2015)	SOP 0.87 / SPP 0.86		CFA Original 2- factors: CFI=0.9, RMSEA=0.09

Table 7. Psychometric Properties

Measure/ reference	Internal consistency	Reliability	Structural validity
CAPS Turkish			
Baş & Siyez (2010)	SOP 0.64 / SPP 0.82		PCA Original 2-factors: 30.5 % variance Proposed 2-factor (18-items): 36.0 % variance
Subsample		SOP 0.63/ SPP 0.72 (2 weeks)	CFA Proposed 2-factor (18-items): CFI=0.90, RMSEA=0.06
CAPS Portuguese			
Bento et al. (2014)	SOP 0.83 / SPP 0.86		PCA: Original 2-factors: 41.4 % variance
Subsample		ICC SOP 0.59 / SPP 0.69 (5 weeks)	
CAPS Spanish			
Castro et al. (2004)			
Sample 1	SOP 0.75 / SPP 0.82		
Sample 2	SOP 0.92 / SPP 0.92		
Subsample		SOP 0.83 / SPP 0.83 (1 week)	

Table 7. Psychometric Properties

Measure/ reference	Internal consistency	Reliability	Structural validity
CAPS Chinese			
Yang et al. (2015)	NR: 4-factor only		CFA Original 2-factors: CFI=0.9, RMSEA=0.055
PSPS-J			
Flett et al. (2012a)	PSP 0.90/ NDP 0.73 NDC 0.57		
Hewitt et al. (2011)			
Sample 1	PSP 0.92/ NDP 0.82 NDC 0.72		
Sample 2	PSP 0.91/ NDP 0.70 NDC0.60		CFA Original 3-factor solution: CFI=0.98, RMSEA= 0.08
Sample 3	PSP 0.89/ NDP 0.78 NDC 0.66		
PCI			
Flett et al. (2012b)			

Table 7. Psychometric Properties

Measure/ reference	Internal consistency	Reliability	Structural validity
Sample 1	0.91		PCA 1-factors: 36.5 % variance
Sample 2	0.91		
FMPS			
Hawkins et al. (2006)	NR: 4-factor only		CFA Original 6-factor solution: 54% of variance CFI=0.92, RMSEA= 0.05 Proposed 4-factor (33-items): 48% of variance CFI=0.90, RMSEA=0.07 (PE & PC merge; CM and DA merge)
Parker & Stumpf (1995)	CM 0.83/ PS 0.74 PE 0.77/ PC 0.78 DA 0.67/ O 0.90		CFA Original 6-factor solution: 64.4 % of variance, GFI=0.86
Siegle & Schuler (2000)	NR: 5-factor only		PCA Proposed 8-factor solution: 59% of variance
Sironic & Reeve (2015)	CM 0.88/ PS 0.85 PE 0.85/ PC 0.82 DA 0.68/ O 0.93		EFA Proposed 4-factor solution: CFI=0.96, RMSEA=0.06 (PE & PC merge; CM and DA merge)

Table 7. Psychometric Properties

Measure/ reference	Internal consistency	Reliability	Structural validity
Stumpf & Parker (2000)	NR		PCA 2-factors: Unhealthy perfectionism 36% of variance, Healthy perfectionism 32% of variance
Taylor et al. (2017)	CM 0.89/ PS 0.69 PE 0.72/ PC 0.69 DA 0.59/ O 0.90		EFA Proposed 1-factor (6-item): CFI=0.96, RMSEA=0.07 (No items from O or CM) Replicated with validation sample
FMPS Chinese			
Chan (2009)	NR: 5-factor only		EFA Original 6-factors: CFI=0.92, RMSEA=0.077 Proposed 5-factor (15-items) 43.1% of variance CFI=0.97, RMSEA=0.056 (D excluded)
Cheng et al. (1999)	NR: 5-factor only		PCA Proposed 5-factors (27-items): 43.1 % of variance
FMPS Portuguese			
Correia et al. (2017)			
Sample 1	NR: 5-factor only		CFA Original 6-factors: CFI=0.81, RMSEA=0.072 Proposed 6-factors (21-items) CFI=0.96, RMSEA=0.045
Sample 2			CFA Proposed 6-factors (21-items), CFI=0.96, RMSEA=0.032

Table 7. Psychometric Properties

Measure/ reference	Internal consistency	Reliability	Structural validity
APS-R			
Sironic & Reeve (2015)	HS 0.88/ D 0.92/ O 0.83		EFA Original 3-factors: CFI=0.96, RMSEA=0.09
Vandiver & Worrell (2002)	HS 0.79/ D 0.87/ O0.79		CFA Original 3-factor: CFI=0.88, RMSEA=0.064
APS-R Spanish			
Sastre-Riba et al. (2016)	HS 0.67/ D 0.85/ O 0.73*		CFA Original 3-factor: CFI 0.85, RMSEA=0.06
APS-R Chinese			
Wang et al. (2009)	NR: 19-items only		EFA Original 3-factors: 50.11% of variance
AMPS			
Rice & Preusser (2002)			
Sample 1	SM 0.91/ CSE 0.86 C 0.87/ NfA 0.85		PCA Original 4 -factors: 52% of variance
Sample 2	SM 0.90/ CSE 0.73 C 0.75/ NfA 0.81		PCA Original 4 -factors: 47% of variance (Small sample)
Rice et al. (2007)	SM=0.84/0.81 (M/F)		PCA Proposed 3 -factors: 53% of variance before rotation,

Table 7. Psychometric Properties

Measure/ reference	Internal consistency	Reliability	Structural validity
	NfA=0.72/0.78 (M/F)		
	C=0.64/0.56 (M/F)		
AMPS Chinese			
Fong & Yuen (2011)	SM 0.47/ CSE 0.65		CFA Original 4-factors: CFI=0.77, RMSEA=0.07
	C 0.69/ NfA 0.83		Proposed 4-factors (23 items): CFI=0.90, RMSEA=0.05

ICC = intraclass correlation coefficient; EFA =Exploratory Factor Analysis; CFA = Confirmatory Factor Analysis; PCA = Principle Component Analysis; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation

3.5. Best Evidence Synthesis

A Best Evidence Synthesis has been summarised in Table 8. Further comments and caveats to the ratings have also been shared below.

Table 8. Best Evidence Synthesis

Measure	Internal consistency	Reliability	Content validity	Structural validity	Hypothesis Testing	Respon-siveness
CAPS	+ + + /?	?	?	- -	++	na
PSPS-J	±	na	?	+ +	+++	na
PCI	++	na	na	?	++	na
FMPS	?	na	?	++/?	+++	na
APS-R	+ + +	na	na	±	++	na
AMPS	++	na	++	?	++	na

Note: +++/--- = Strong evidence of positive/negative finding; ++/-- = Moderate evidence of positive/negative finding; +/- = Limited evidence of positive/negative finding; ± = Conflicting findings; ? Indeterminate finding; na = not assessed

Internal Consistency

CAPS: Criteria was met in two good quality studies, suggesting evidence in this domain should be rated as strong. However, following indeterminate findings by McCreary et al. (2004), and in light of poor evidence of structural validity, queries have been indicated with a question mark.

PSPS-J: Criteria was met by one good quality study, but not found for the Nondisclosure of Perfectionism subscale in the second study sample. Evidence is therefore conflicting.

PCI: Criteria was met by one good quality study. Evidence of internal consistency was therefore rated as moderate.

FMPS: Indeterminate findings were reported across three good quality studies. The Doubts about Actions subscale consistently fell short of the criteria. Evidence was therefore rated as indeterminate.

APS-R: Criteria was met by two good quality studies. Evidence was therefore rated as strong.

AMPS: Criteria was met by one good quality study. Evidence was therefore rated as moderate. In a second study by the same authors, the Compulsiveness subscale did not meet the criteria. As this study only assessed three out of the four originally proposed subscales, hence this study was not included in the BES, but should be noted as a caveat.

Reliability

CAPS: Reliability was assessed by one good quality study. For the total sample, criteria could not be met. Authors reported higher validity in a small subsample of older children, hence evidence was rated as indeterminate.

Content validity

AMPS: One good quality study assessed content validity by consulting the target population. Measures were adjusted accordingly. Evidence was therefore rated as moderate.

For all other instruments, content validity was either not assessed, or had only been assessed by studies which were of poor quality (rated as indeterminate).

Structural Validity

CAPS: The criteria was only met by one good quality study, but was in other studies suggestive of a poor or equivocal fit. Findings indicated moderate evidence of poor structural validity of the original two-factor structure of the CAPS.

PSPS-J: Criteria was met by one good quality study. Evidence or structural validity was therefore rated as moderate.

PCI: Structural validity was only explored by using Principle Component Analysis. Evidence was therefore rated as indeterminate.

FMPS: The criteria was met by two good quality studies. Inconsistencies were however reported across studies, with some suggesting a reduction of items and subscales would produce a better fit. Evidence was therefore rated as conflicting.

APS-R: Findings were conflicting across two good quality studies.

AMPS: Evidence rated as indeterminate, as Principle Component Analysis was used.

Hypothesis Testing

For studies where hypothesis were stated, scores on the instrument tended to be consistent with the predictions made. The evidence was therefore consistently rated as moderate or strong across the included studies.

4. Discussion

The aim of the present review was to systematically examine and evaluate the evidence of psychometric properties, reported for measures of perfectionism used with children and adolescents. A critical appraisal tool, the COSMIN checklist, was used to rank the methodological rigor of the studies included in the review. The reported psychometric properties were thereafter compared against predefined criteria to determine if the validity and reliability reported for each measure could be considered adequate. A Best Evidence Synthesis approach was used to summarise the findings of studies which had assessed original versions of the instruments (see Table 8).

A comprehensive search strategy was used to find 2538 records, of which 26 articles met the study's inclusion criteria. Collectively, they assessed psychometric domains of six self-report instruments. The most commonly evaluated questionnaire was the FMPS, followed by the CAPS. For the other instruments, studies were few, and were often conducted by the instrument developers.

In assessing the methodological quality, the COSMIN ratings were relaxed so that studies could be rated as 'good', even if they had not reported how missing items were treated. Initially, the rating manual was followed, but after discovering that almost all of the articles included would have had an upper limit of a 'fair' rating, this standard was changed *a posteriori*. Following this adjustment, a broader variation in ratings across the different domains could be observed. Most studies assessed internal consistency and structural validity, and the methodological quality was typically rated as 'good'. For other properties, including test-retest reliability, hypothesis testing and cross-cultural validity, ratings were typically lower because information was absent or unclear. This observation highlights a predicament; while the rigor of the COSMIN could be seen as a limitation as

it may lead to the overall quality of studies being underrepresented, there is also a need for publications to become more transparent about the methodology used. Studies could for example be supplemented with appendices in electronic versions of journals.

There was a concerning lack of information about some questionnaire properties. No studies presented information about measurement error, few studies assessed test-retest reliability or content validity, and in no study was responsiveness explored. These are all important aspects which will require investigation before a fully informed decision of which instrument to use can be made. The absence of knowledge about test-retest reliability is perhaps the most urgent to address. Test-retest reliability will be important as there is an increased drive for prospective studies and clinical interventions of perfectionism (Morris & Lomax, 2014). An estimation of the consistency in scores over time will be needed before one can conclude that changes in scores were achieved as a function of another variable, such as stressful life events or an intervention.

Another limitation across studies was the lack of clarity over the practical elements of administering and scoring the instruments, for example the completion time, suggested cut-off points, and T-scores to enable the comparison of a person's individual score against norms from an equivalent age and gender group. These aspects would help to increase the interpretability of scores and could guide the choice of instruments.

Based on psychometric information alone, there was no one measure which demonstrated superiority over the other measures included for review. The three measures which had been most researched, and which received the highest BES ratings of evidence, were the CAPS, FMPS, and APS-R. There were however some individual concerns about each instrument. Whilst the APS-R is emerging as a promising measure, it had only been evaluated by two good quality studies. Evidence of good internal consistency was reported, but the studies demonstrated inconsistent findings for structural validity, which warrant some concerns. The FMPS had some evidence of structural validity, but many of the studies reported a preference for a reduced version of the instrument with fewer subscales. Furthermore, the Doubts about Actions subscale of this measure did not demonstrate adequate internal consistency. Studies assessing the CAPS found evidence of adequate internal consistency, but not of structural validity, which may indicate that the self-oriented subscale would be better conceptualised as

two subscales. While taking these concerns into consideration, the CAPS was the only instrument out of these three measures which had been specifically adapted for the use by children and adolescents, and age-appropriate norms had also been published. The use of the CAPS could therefore be preferred, especially when used with younger populations.

There were some limitations to this review. Firstly, the search, screening, and selection of articles were all done by the same researcher (SM). To account for possible errors or biases, the screening of titles and abstracts of articles was done twice. Nevertheless, the study would have benefitted from the verification of findings by another researcher. Secondly, grey literature was not included in this review. A primary search was conducted, which indicated that no articles could be found. Nevertheless, if this study was replicated at a later stage, this should be taken into consideration, as the gap in reported properties such as test-retest reliability may reflect a possible publication bias. Furthermore, it was beyond the scope of the present review to include all articles which had assessed any psychometric properties of a measure, for example reporting internal consistency as part of a study. The inclusion criteria were made stringent so that articles would be selected in a systematic manner. This may however have led to some relevant studies being missed. In contrast, Leone and Wade (2017) used a more inclusive approach and reported data from a large number of studies (n=76). How their findings compared to the present review was not assessed until after this study had been completed. Despite using different methodologies, both reviews were in agreement with the conclusions made about most psychometric properties. The overlap in findings suggest that the caveats mentioned above are due to a gap in research about psychometric properties of these instruments, rather than as an effect of important articles being missed at the screening stage. A difference was found in the assessment of test-retest reliability. Leone and Wade (2017) had used a more lenient criterion, and concluded that the test-retest reliability of the CAPS was adequate, whereas in this study evidence of test-retest validity of the CAPS was rated as indeterminate.

In addition to the need to evaluate neglected measurement properties and provide clearer instructions about the administration and scoring of instruments, future research should focus on exploring how young people relate to the perfectionism construct. If perfectionism is perceived as positive, young people may for example overreport tendencies to hold perfectionist ideals. The importance of considering cultural differences

has also been highlighted (Yang et al., 2015; Cheng, 1999; Chan, 2009). It is important to further investigate construct validity to assist in the interpretation of scores, and to better understand the structural validity of the measures. A last recommendation includes for researchers to aim for smaller differences in ages between participants, or to recruit large enough sample sizes to control for age, or to split participants into subsamples. This is because some psychometric properties have been suggested to change with age.

In conclusion, this review has identified that many perfectionism measures used with children and adolescents lack, in part, evidence of adequate psychometric properties. This highlights a need for further exploration of the reliability and validity of instruments, and guidance of how scores should be interpreted. A clear recommendation about which instrument to use cannot, at this point in time, be made. A tentative recommendation would however be to use the CAPS, as it was one of the most well-documented measures, it had been specifically adapted for the use by children and adolescents, and age-appropriate norms were available.

References

Adler, A. (1964) *Striving for superiority.* In H. L. Ansbacher & R. Ansbacher Eds.), *The individual psychology of Alfred Adler: A systematic presentation in selections from his writings* (pp. 101–125). New York: Harper & Row.

Baş, A. U., & Siyez, D. M. (2010) Adaptation of the Child and Adolescent Perfectionism Scale to Turkish: The validity and reliability study. *İlköğretim Online*, *9,* (3).

Bento, C., Pereira, A. T., Saraiva, J. M., & Macedo, A. (2014) Children and adolescent perfectionism scale: validation in a Portuguese adolescent sample. *Psicologia: Reflexão e Crítica*, *27,* (2), 228-232.

Bieling, P. J., Israeli, A., Smith, J., & Antony, M. M. (2003) Making the grade: The behavioural consequences of perfectionism in the classroom. *Personality and Individual Differences*, *35,* (1), 163-178.

Castro, J., Gila, A., Gual, P., Lahortiga, F., Saura, B., & Toro, J. (2004) Perfectionism dimensions in children and adolescents with anorexia nervosa. *Journal of adolescent health,* 35, (5), 392-398.

Chan, D. W. (2009) Dimensionality and typology of perfectionism: The use of the Frost Multidimensional Perfectionism Scale with Chinese gifted students in Hong Kong. *Gifted Child Quarterly*, *53,* (3), 174-187.

Chang, E. C. (2006) Perfectionism and Dimensions of Psychological Well–Being in a college Student Sample: A test of a stress–Mediation Model. *Journal of Social and Clinical Psychology*, *25,* (9), 1001-1022.

Cheng, S. K., Chong, G. H., & Wong, C. W. (1999) Chinese frost multidimensional perfectionism scale: A validation and prediction of self-esteem and psychological distress. *Journal of Clinical Psychology*, *55,* (9), 1051-1061.

Correia, M., Rosado, A., & Serpa, S. (2017) Psychometric properties of the Portuguese version of the Frost Multidimensional Perfectionism Scale. *International Journal of Psychological Research*, *10,* (1), 8-17.

Dour, H. J., & Theran, S. A. (2011) The interaction between the superhero ideal and maladaptive perfectionism as predictors of unhealthy eating attitudes and body esteem. *Body image*, *8,* (1), 93-96.

Fichman, L., Koestner, R., & Zuroff, D. C. (1994) Depressive styles in adolescence: Assessment, relation to social functioning, and developmental trends. *Journal of Youth and Adolescence*, *23*, (3), 315-330.

Flett, G. L., Coulter, L. M., & Hewitt, P. L. (2012a) The Perfectionistic Self-Presentation Scale–Junior Form: Psychometric properties and association with social anxiety in early adolescents. *Canadian Journal of School Psychology*, *27*, (2), 136-149.

Flett, G. L., & Hewitt, P. L. (2013) Disguised distress in children and adolescents "flying under the radar" why psychological problems are underestimated and how schools must respond. *Canadian Journal of School Psychology*, *28*, (1), 12-27.

Flett, G. L., Hewitt, P. L., Besser, A., Su, C., Vaillancourt, T., Boucher, D., ... & Gale, O. (2016) The Child–Adolescent Perfectionism Scale: Development, psychometric properties, and associations with stress, distress, and psychiatric symptoms. *Journal of Psychoeducational Assessment*, *34*, (7), 634-652.

Flett, G. L., Hewitt, P. L., Blankstein, K. R., & Gray, L. (1998) Psychological distress and the frequency of perfectionistic thinking. *Journal of Personality and Social Psychology*, 75, (5)

Flett, G. L., Hewitt, P. L., Boucher, D. J., Davidson, L. A., & Munro, Y. (2000) The child-adolescent perfectionism scale: Development, validation, and association with adjustment. *Unpublished manuscript*.

Flett, G. L., Hewitt, P. L., Demerjian, A., Sturman, E. D., Sherry, S. B., & Cheng, W. (2012b) Perfectionistic automatic thoughts and psychological distress in adolescents: An analysis of the Perfectionism Cognitions Inventory. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *30*, (2), 91-104.

Fong, R. W., & Yuen, M. (2011) Perfectionism in Chinese elementary school students: Validation of the Chinese adaptive/maladaptive perfectionism scale. *Talent development and excellence*.

Frost, R. O., Marten, P., Lahart, C., & Rosenblate, R. (1990) The dimensions of perfectionism. *Cognitive therapy and research*, *14*, (5), 449-468.

Hawkins, C. C., Watt, H. M., & Sinclair, K. E. (2006) Psychometric properties of the Frost Multidimensional Perfectionism Scale with Australian adolescent girls: Clarification of multidimensionality and perfectionist typology. *Educational and psychological measurement*, *66*, (6), 1001-1022.

Hill, A. P., Appleton, P. R., & Mallinson, S. H. (2016) Development and initial validation of the Performance Perfectionism Scale for Sport (PPS-S). *Journal of Psychoeducational Assessment*, *34*, (7), 653-669.

Hewitt, P. L., Blasberg, J. S., Flett, G. L., Besser, A., Sherry, S. B., Caelian, C., ... & Birch, S. (2011) Perfectionistic self-presentation in children and adolescents: Development and validation of the Perfectionistic Self-Presentation Scale—Junior Form. *Psychological assessment*, *23*, (1), 125.

Hewitt, P. L., Caelian, C. F., Flett, G. L., Sherry, S. B., Collins, L., & Flynn, C. A. (2002) Perfectionism in children: Associations with depression, anxiety, and anger. *Personality and Individual Differences*, *32*, (6), 1049-1061.

Hewitt, P. L., & Flett, G. L. (1991) Perfectionism in the self and social contexts: conceptualization, assessment, and association with psychopathology. *Journal of personality and social psychology*, *60*, (3), 456.

Hu, L. T., & Bentler, P. M. (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*, (1), 1-55.

Huggins, L., Davis, M. C., Rooney, R., & Kane, R. (2008) Socially prescribed and self-oriented perfectionism as predictors of depressive diagnosis in preadolescents. *Journal of Psychologists and Counsellors in Schools*, *18*, (2), 182-194.

Jessor, R., Donovan, J. E., & Costa, F. (1990). Personality, perceived life chances, and adolescent health behavior. In K. Hurrelmann & F. Lösel (Eds.), *Health hazards in adolescence*, (pp. 25–42). New York: Walter de Gruyter.

Jones, J., Ramirez, R. R., Davies, M., Canino, G., & Goodwin, R. D. (2008) Suicidal behaviors among adolescents in Puerto Rico: Rates and correlates in clinical and community samples. *Journal of Clinical Child & Adolescent Psychology*, *37*, (2), 448-455.

Leone, E. M., & Wade, T. D. (2017) Measuring perfectionism in children: a systematic review of the mental health literature. *European child & adolescent psychiatry*, 1-15.

McCreary, B. T., Joiner, T. E., Schmidt, N. B., & Ialongo, N. S. (2004) The structure and correlates of perfectionism in African American children. *Journal of Clinical Child and Adolescent Psychology*, *33*, (2), 313-324.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, *6*, (7).

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso. J., Stratford, P.W., Knol, D.L., Bouter, L.M., de Vet, H.C.W (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*, 19, 539-549.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... & de Vet, H. C. (2012) COSMIN checklist manual. *Amsterdam: COSMIN*.

Morris, L., & Lomax, C. (2014) Assessment, development, and treatment of childhood perfectionism: A systematic review. *Child and Adolescent Mental Health*, *19*, (4), 225-234.

Nock, M. K., & Prinstein, M. J. (2005) Contextual features and behavioral functions of self-mutilation among adolescents. *Journal of abnormal psychology*, *114*, (1), 140.

O'Connor, R. C., Rasmussen, S., & Hawton, K. (2010) Predicting depression, anxiety and self-harm in adolescents: The role of perfectionism and acute life stress. *Behaviour research and therapy*, *48*, (1), 52-59.

O'Connor, R. C., Dixon, D., & Rasmussen, S. (2009) The structure and temporal stability of the Child and Adolescent Perfectionism Scale. *Psychological Assessment*, *21*, (3), 437.

Park, T., Reilly-Spong, M., & Gross, C. R. (2013) Mindfulness: a systematic review of instruments to measure an emergent patient-reported outcome (PRO). *Quality of Life Research*, *22*, (10), 2639-2659.

Parker, W. D., & Stumpf, H. (1995) An examination of the Multidimensional Perfectionism Scale with a sample of academically talented children. *Journal of Psychoeducational Assessment*, *13*, (4), 372-383.

Piers, E. V. (1969) Manual for the Piers-Harris Children's Self Concept Scale:(the Way I Feel about Myself). *Counselor Recordings and Tests*

Rice, K. G., & Preusser, K. J. (2002) The adaptive/maladaptive perfectionism scale. *Measurement and evaluation in Counseling and Development*, *34*, (4), 210

Rice, K. G., Kubal, A. E., & Preusser, K. J. (2004) Perfectionism and children's self-concept: Further validation of the Adaptive/Maladaptive Perfectionism Scale. *Psychology in the Schools*, *41*, (3), 279-290.

Rice, K. G., Leever, B. A., Noggle, C. A., & Lapsley, D. K. (2007) Perfectionism and depressive symptoms in early adolescence. *Psychology in the Schools*, *44*, (2), 139-156.

Sastre-Riba, S., Pérez-Albéniz, A., & Fonseca-Pedrero, E. (2016) Assessing perfectionism in children and adolescents: Psychometric properties of the Almost Perfect Scale Revised. *Learning and Individual Differences*, 49, 386-392.

Schellingerhout, J. M., Verhagen, A. P., Heymans, M. W., Koes, B. W., Henrica, C., & Terwee, C. B. (2012) Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review. *Quality of Life Research*, *21*, (4), 659-670.

Shafran, R., & Mansell, W. (2001) Perfectionism and psychopathology: A review of research and treatment. *Clinical Psychology Review*, *21*, (6), 879-906.

Siegle, D., & Schuler, P. A. (2000). Perfectionism differences in gifted middle school students. *Roeper Review*, *23*, (1), 39-44.

Sironic, A., & Reeve, R. A. (2015) A combined analysis of the Frost Multidimensional Perfectionism Scale (FMPS), Child and Adolescent Perfectionism Scale (CAPS), and Almost Perfect Scale—Revised (APS-R): Different perfectionist profiles in adolescent high school students. *Psychological assessment*, *27*, (4), 1471.

Slaney, R. B., & Ashby, J. S. (1996) Perfectionists: Study of a criterion group. *Journal of Counseling & Development*, 74, (4), 393-398.

Slaney, R. B., Rice, K. G., Mobley, M., Trippi, J., & Ashby, J. S. (2001) The revised almost perfect scale. *Measurement and evaluation in counseling and development*, *34*, (3), 130.

Stoeber, J., & Otto, K. (2006) Positive conceptions of perfectionism: Approaches, evidence, challenges. *Personality and social psychology review*, *10*, (4), 295-319.

Stornelli, D., Flett, G. L., & Hewitt, P. L. (2009) Perfectionism, achievement, and affect in children: A comparison of students from gifted, arts, and regular programs. *Canadian Journal of School Psychology*, *24*, (4), 267-283.

Stumpf, H., & Parker, W. D. (2000) A hierarchical structural analysis of perfectionism and its relation to other personality characteristics. *Personality and individual differences*, *28*, (5), 837-852.

Taylor, E. P., Couper, R., & Butler, C. M. (2017) Adolescent perfectionism: Structural features of the Frost Multidimensional Perfectionism Scale and correlates with attachment and psychopathology. *Psychology and Psychotherapy: Theory, Research and Practice*, *90*, (4), 686-704.

Van Tulder, M., Furlan, A., Bombardier, C., Bouter, L., & Editorial Board of the Cochrane Collaboration Back Review Group (2003) Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine*, *28*, (12), 1290-1299.

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., ... & de Vet, H. C. (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of clinical epidemiology*, *60*, (1), 34-42.

Terwee, C.B., Mokkink, L.B., Knol, D.L., Ostelo, R.W.J.G., Bouter, L.M., & de Vet, H.C.W. (2011) Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research.*

Terwee, C. B., de Vet, H.C.W., CAC, P., & Mokkink, L. B. (2011) Protocol for systematic reviews of measurement properties. *COSMIN: Knowledgecenter Measurement Instruments*.

Vandiver, B. J., & Worrell, F. C. (2002) The reliability and validity of scores on the Almost Perfect Scale–Revised with academically talented middle school students. *Journal of Secondary Gifted Education*, *13*, (3), 108-119.

Wang, K. T., Yuen, M., & Slaney, R. B. (2009) Perfectionism, depression, loneliness, and life satisfaction: A study of high school students in Hong Kong. *The Counseling Psychologist*, *37*, (2), 249-274.

Weissman, A. N., & Beck, A. T. (1978) Development and validation of the Dysfunctional Attitude Scale: A preliminary investigation.

Yang, H., Hong, C., Tao, X., & Zhu, L. (2015) Revising the child and adolescent perfectionism scale: a test of the four-factor structure in a Chinese sample. *Measurement and Evaluation in Counseling and Development, 48,* (3), 192-203.

Appendix 1. PsychInfo Search Terms

1	**Perfection*.mp.**
2	child*.mp.
3	adolescen*.mp.
4	teen*.mp.
5	youth.mp.
6	young.mp.
7	measure*.mp.
8	assess*.mp.
9	scale.mp.
10	2 or 3 or 4 or 5 or 6
11	7 or 8 or 9
12	1 and 10 and 11